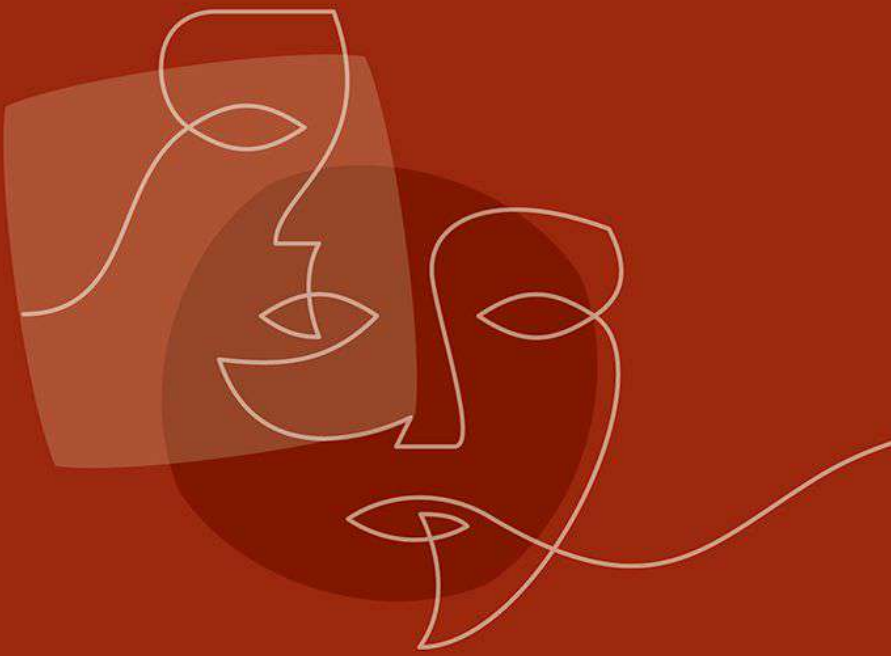
OXFORD

SELF AND IDENTITY



TRENTON MERRICKS

Self and Identity

Self and Identity

TRENTON MERRICKS

CLARENDON PRESS • OXFORD

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© Trenton Merricks 2022

The moral rights of the author have been asserted

First Edition published in 2022
Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in
a retrieval system, or transmitted, in any form or by any means, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2021944981

ISBN 978–0–19–284343–2

DOI: 10.1093/oso/9780192843432.001.0001

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

For Laura

Contents

Introduction 1

1. What Matters in Survival 7
   - I. Appropriate First-Personal Anticipation and Appropriate Future-Directed Self-Interested Concern 7
   - II. My Answer to the What Question 14
   - III. Consciousness and Survival 20
   - IV. What Matters to You with Regard to the Future 23
   - V. Conclusion 28

2. On the Sufficiency of Personal Identity 29
   - I. My Answer to the Why Question 31
   - II. More on the Metaphysics of Persistence 35
   - III. Not the Criterion of Personal Identity over Time 46
   - IV. An Unanswered Question 54
   - V. Conclusion 55

3. On the Necessity of Personal Identity 57
   - I. An Argument for the Necessity of Personal Identity 57
   - II. Parfit's Argument against the Necessity of Personal Identity 65
   - III. Parfit's Argument, Stage Theory, and Perdurance 71
   - IV. Parfit's Argument and Endurance 75
   - V. Psychological Connectedness and Psychological Continuity 83
   - VI. Conclusion 85

4. The Same Self 87
   - I. Three Selfers 87
   - II. First-Personal Access to a Point of View 94
   - III. The Same Self and Numerical Identity 101
   - IV. Growing Up 103
   - V. Other Transformations 107
   - VI. Conclusion 111

5. The Same Self-Narrative 113
I. The Self-Narrative Account 113
II. The Same Self-Narrative and the Same Self 118
III. The Same Self-Narrative and Numerical Identity 120
IV. Growing Up Redux 121
V. Other Changes in Self-Narrative 126
VI. Other Work for Self-Narrative and the Same Self 128
VII. Conclusion 131

6. Agential Continuity and Narrative Continuity 133
I. The Agential Continuity Account 133
II. The Narrative Continuity Account 139
III. Agential Continuity, Narrative Continuity, and Numerical Identity 142
IV. Some Significant Transformations 145
V. Other Work for Agential Continuity and Narrative Continuity 150
VI. More on Psychological Connectedness and Psychological Continuity 151
VII. Conclusion 155

7. The Hope of Glory 157
I. The Hope of Survival 157
II. The Hope of Transformation 160
III. The Hope and Psychological Continuity 163
IV. Survival Does Not Come in Degrees 164
V. The Tedium Objection 167
VI. The Irrationality Objection 171
VII. Conclusion 174

References 177
Index 183

Acknowledgments

I presented parts of this book as talks at the APA Pacific Pre-Conference on Themes in Transformative Experience, the Ranch Metaphysics Workshop, the University of Virginia, Virginia Commonwealth University, the New England Workshop in Metaphysics, Simon Fraser University, Wake Forest University, and Syracuse University. Thanks to all who attended those talks, and in particular to those who pointed out problems or made helpful suggestions, especially Shamik Dasgupta, Naomi Dershowitz, Katie Elliot, Mark Heller, Dave Ingram, Gene Mills, Connie Rosati, Erica Shumener, Donald Smith, and Jennifer Wang.

Several groups read a draft of the book and then helped me to improve it, including a seminar taught by Mike Rea at Notre Dame; a seminar taught by Dan Korman at UC Santa Barbara; the Paul Lab at Yale, run by Laurie Paul; and Yale's ELLMM City Reading Group. A reading group here at the University of Virginia discussed each chapter with me in detail, over the course of a semester; I am particularly indebted to the members of this group, especially Kirra Hyde and Bill Vincent.

Tal Brewer, Jim Cargile, Rebecca Chan, Matt Duncan, Harold Langsam, Laurie Paul, Jack Spencer, and Eleonore Stump helped me with portions of this book. Elizabeth Barnes, Mike Bergmann, Ross Cameron, Jim Darcy, Zac Irving, Dan Korman, Mark Murphy, Eric Olson, Mike Rea, Bradley Rettler, Rebecca Stangl, and two readers for Oxford University Press gave me insightful feedback on the entire manuscript. I thank them all for being so generous with their time and for making this book much better than it would otherwise have been. Thanks also to Greg Breeding for designing the cover and doing the artwork.

T. M.
Charlottesville, VA

Introduction

This book revolves around the following two questions:

> *The What Question*: What is it for a person at a future time to have (at that time) what matters in survival for you?
>
> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

The What Question is distinct from the Why Question. But it would be understandable if someone conflated them. After all, the answers to both can start with: 'a person at a future time will have (at that time) what matters in survival for you because . . . '.

Again, the What Question is distinct from the Why Question. One reason that I say this begins with my answer to the What Question, which is defended in Chapter 1: its being appropriate for you to first-personally anticipate the experiences that that person will have at that future time; and if that person will have good (or bad) experiences at that future time, its being appropriate for you to have future-directed self-interested concern with regard to those experiences.

Now ask:

> *The Wordier Why Question*: What way of being related to a (conscious) person at a future time explains why it is appropriate for you to both (i) first-personally anticipate the experiences that that person will have at that time and also (ii) have future-directed self-interested concern with regard to the good (or bad) experiences that that person will have at that time?

The answer to the Wordier Why Question is surely not (anything like) my answer to the What Question.

Perhaps you need convincing. Then assume for reductio that the answer to the Wordier Why Question is my answer to the What

Self and Identity. Trenton Merricks, Oxford University Press. © Trenton Merricks 2022.
DOI: 10.1093/oso/9780192843432.003.0001

Question. So its being appropriate for you to first-personally anticipate, and have future-directed self-interested concern with regard to, the experiences a person will have at a future time (is a way of your being related to that person that) explains why it is appropriate for you to first-personally anticipate, and have future-directed self-interested concern with regard to, those experiences. In other words, the relevant anticipation and concern being appropriate explains why that anticipation and concern are appropriate. But that cannot be right.

Suppose that my answer to the What Question is correct. Then the Wordier Why Question 'unpacks' the Why Question. More to the point, then the Wordier Why Question and the Why Question have the same answer. We have already seen that—given my answer to the What Question—the What Question and the Wordier Why Question do not have the same answer. So—given my answer to the What Question—we should conclude that the What Question and the Why Question do not have the same answer. So my answer to the What Question leads me to conclude that the What Question and the Why Question are distinct questions.

My answer to the What Question is not too controversial. As we shall see in Chapter 1, my answer coheres with what a lot of philosophers already say. That is, a lot of philosophers answer the What Question at least in part in terms of appropriate first-personal anticipation or appropriate future-directed self-interested concern. But this is not how those philosophers answer the Why Question. Instead, most of those philosophers answer the Why Question in terms of some sort of psychological connectedness or psychological continuity. So I am not alone in taking the What Question and the Why Question to have distinct answers. So I am not alone in taking them to be distinct questions.

My answer to the What Question is in terms of appropriate first-personal anticipation and appropriate future-directed self-interested concern. So I deny that the answer to the What Question is in terms of identity. Again, I deny that your being identical with a person at a future time is what it is for that person, at that time, to have what matters in survival for you. So I shall say that identity is not what matters in survival. But this does not imply that identity is irrelevant to what matters in survival. In particular, this does not imply that identity is irrelevant to the answer to the Why Question. For the answer to the Why Question is distinct from the answer to the What Question.

Identity is not what matters in survival. But identity delivers what matters in survival. This is because—so I say—the answer to the Why Question is numerical identity. That is, I say that your being numerically identical with a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you.

My answer to Why Question might seem obvious. For it might seem obvious that it is appropriate for you to first-personally anticipate, and have future-directed self-interested concern with regard to, a future experience because you yourself will have that experience. But it turns out that, unlike my answer to the What Question, my answer to the Why Question is quite controversial.

As we shall see in Chapter 2, the claim that numerical identity is a good answer to the Why Question is widely denied in the personal identity literature that is concerned primarily with what matters in survival. In that literature—and as noted above—the Why Question is typically answered in terms of some sort of psychological connectedness or psychological continuity.

Moreover, the claim that numerical identity is a good answer to the Why Question is not really defended in the personal identity literature that is primarily concerned with the metaphysics of persistence. This is unsurprising. For the Why Question is not a question about the metaphysics of persistence, not even as it applies to persons in particular. Rather, the Why Question is a question about what matters in survival.

The defense of my answer to the Why Question presented in Chapter 2 relies on a particular metaphysics of persistence, namely, that persons persist by 'enduring'. Those who reject this metaphysics of persistence can resist my defense of my answer to the Why Question. Chapters 2 and 3 will consider some options for non-endurantists with regard to answering the Why Question. But I do think that persons endure. So I do think that numerical identity is a good answer to the Why Question.

Chapter 2 defends the conclusion that numerical identity is a good answer to the Why Question. So if a (conscious) person at a future time is numerically identical with you, then that person will have, at that time, what matters in survival for you. So—at least if persons endure—personal identity over time is sufficient for what matters in survival. Chapter 3 defends the further conclusion that personal identity over time is necessary for what matters in survival. Chapter 3 also

shows that Derek Parfit's argument against the necessity of personal identity for what matters in survival fails.

Again, Chapter 3 defends the conclusion that a person at a future time will have (at that time) what matters in survival for you only if you will be the same person as that person. I think that you endure. So I conclude that a person at a future time will have (at that time) what matters in survival for you only if you will be numerically identical with that person. So I conclude that not only is numerical identity a good answer to the Why Question, but also that every good answer implies numerical identity.

Chapter 4 considers the view that every good answer to the Why Question implies having the 'same self', where this is understood as being alike with regard to the values, desires, and projects that 'make you the person you are'. Chapter 5 focuses on an answer to the Why Question in terms of being alike with regard to having the same 'self-narrative', that is, the same story of a life.

Some will take the answers to the Why Question considered in Chapters 4 and 5 to be, at bottom, the same answer. For they will say that the values, desires, and projects that 'make you the person you are' just are the values, desires, and projects that are included in your self-narrative. But others will deny this. These others will take answers to the Why Question that imply having the 'same self' to be independent of any claim about self-narrative.

But all can agree that the answers to the Why Question considered in Chapters 4 and 5 are alike in a crucial way: each of those answers implies being psychologically connected, and, in particular, being psychologically alike. That is, each of those answers implies that a person at a future time will have (at that time) what matters in survival for you only if the way you are now is psychologically like the way that person will be at that time in some substantive way, such as having the same values, desires, and projects, or having the same self-narrative.

Consider those transformations that involve a change in self-narrative and that involve a change in self, that is, a change in the relevant values, desires, and projects. I defend the claim that such a transformation would be good for you only if the resulting transformed person would have what matters in survival for you. I also motivate the claim that some such transformations would be good for you, or at least good for someone who is currently quite bad, or currently quite

young. All this gives us one reason (among others) to reject the answers to the Why Question considered in Chapters 4 and 5.

Chapter 6 considers an answer to the Why Question that—like the answer considered in Chapter 5—is in terms of narrative. But unlike the answer considered in Chapter 5, the answer considered in Chapter 6 is in terms of narrative continuity, as opposed to narrative connectedness. In particular, this answer is in terms of overlapping local narratives, where a 'local narrative' is a narrative that characterizes a single action. Another answer to the Why Question considered in Chapter 6 is in terms of agential continuity. Agential continuity is constituted by overlapping instances of 'agential connectedness', that is, overlapping instances of a person's choosing to act that results in that person's having various psychological states.

The answers to the Why Question considered in Chapter 6 have an advantage over the answers to the Why Question considered in Chapters 4 and 5. The answers to the Why Question considered in Chapter 6 are consistent with a person at a future time having (at that time) what matters in survival for you even if the way you are now is not at all psychologically like the way that person will be at that time. So these answers can accommodate transformations being good for you, at least if those transformations are gradual enough to preserve agential and narrative continuity.

But we should still reject the answers to the Why Question considered in Chapter 6. One reason (among others) for rejecting those answers begins by supposing that you will become an evil person, but not as a result of your own actions. That is, you will be turned into an evil person. This would be bad for you. And I argue that this would be bad for you in a way other than the way that ceasing to exist would be bad for you. But we shall see that the answers considered in Chapter 6 imply that this would be bad for you only in the way that ceasing to exist would be bad for you.

Chapter 6 concludes by showing how the problems with the specific answers to the Why Question considered in Chapters 4, 5, and 6 should lead us to deny that every good answer to the Why Question must be in terms of psychological connectedness or psychological continuity. So the point of Chapters 4–6 is not merely to oppose a handful of answers to the Why Question that are not consistent with my answer, which is in terms of numerical identity. Rather, the point

of those chapters is to block a likely objection to my answer, an objection that is independent of the specific answers opposed in Chapters 4–6. That objection is that my answer must be wrong—and so my defense of my answer must somehow fail—because the answer to the Why Question must be in terms of some sort of psychological connectedness or psychological continuity.

Chapter 7 is the last chapter and turns to Last Things. I say that the idea of personal immortality is the idea that there will always be someone who will have what matters in survival for you. So—given how I answer the What Question—the idea of personal immortality is not the idea that there will always be someone (conscious) who is identical with you. Nevertheless—given what I say about the Why Question—you will enjoy personal immortality if and only if there will always be someone (conscious) who is identical with you. This chapter then goes on to show how the views defended in the preceding chapters allow us to respond to familiar objections to immortality's possibility and desirability.

1

What Matters in Survival

Consider:

> *The What Question*: What is it for a person at a future time to have (at that time) what matters in survival for you?

I begin this chapter by clarifying the ideas that are invoked in my answer to the What Question. Then I motivate my answer, which is: its being appropriate for you to first-personally anticipate the experiences that that person will have at that future time; and if that person will have good (or bad) experiences at that future time, its being appropriate for you to have future-directed self-interested concern with regard to those experiences. This chapter also distinguishes the What Question from other questions with which it might be conflated.

I. Appropriate First-Personal Anticipation and Appropriate Future-Directed Self-Interested Concern

G. W. Leibniz greatly admired Chinese culture and philosophy (see, e.g. Perkins, 2004). He surely would have thought that there are reasons to desire becoming the ruler of China—but only under certain conditions. Leibniz says:

> Suppose that some person all of a sudden becomes the king of China, but only on the condition that he forgets what he has been, as if he were born anew; practically . . . wouldn't that be the same as if he were annihilated and a king

Self and Identity. Trenton Merricks, Oxford University Press. © Trenton Merricks 2022.
DOI: 10.1093/oso/9780192843432.003.0002

> of China created at the same instant in his place? That is something this individual would have no reason to desire.
>
> ('Discourse on Metaphysics', §34 [1989, 66]; see also Leibniz, 'Letter to Molanus' [1989, 243])

Suppose that you will forget all that you have been and then become the king of China (or the queen of North America, or something else really impressive). Then Leibniz would say that this is the practical equivalent of your being annihilated and replaced by a king (or a queen or . . .). And so Leibniz would insist that your first-personally anticipating, or having self-interested concern with regard to, the king's post-amnesia experiences would be mistaken, or confused, or wrong, or—as I shall put it—not appropriate.

Of course, Leibniz thinks that sometimes your first-personally anticipating, and having self-interested concern with regard to, future experiences is appropriate. This is appropriate, according to Leibniz, just in case you will remember, when having those future experiences, what you have been. The passage from Leibniz quoted above is meant to illustrate this very point. For that passage is immediately preceded by:

> Thus the immortality required in morality and religion does not consist merely in this perpetual subsistence common to all substances, for without the memory of what one has been, there would be nothing desirable about it.
>
> ('Discourse on Metaphysics', §34 [1989, 66])

Remembering an experience involves looking back at that experience. Remembering an experience is not the same thing as believing that you had that experience. For it is possible for you to believe that you had a past experience, but not to remember that experience. For example, you might not remember your first birthday party at all, but still believe that you ate cake at it, and believe this because you have seen a video of that party.

Remembering an experience involves looking back at a past experience. You can also look ahead to a future experience. This is first-personally anticipating an experience. First-personally anticipating an experience is not the same thing as believing that you will have that experience. For I think that it is possible for you to believe that you will have a future experience but not first-personally anticipate that experience.

Leibniz should agree that this is possible. For suppose that Leibniz believes that he will lose all his memories, and then become the king of

China and have various positive experiences. Then Leibniz believes that he will have those experiences. But Leibniz thinks that it is not appropriate for him to first-personally anticipate those experiences. Suppose that Leibniz stays true to his convictions. Then he will not first-personally anticipate those experiences. But, again, Leibniz believes that he will have those experiences.[1]

Let *future-directed* self-interested concern be self-interested concern with regard to what will occur at a future time. I think that first-personally anticipating an experience is more like having future-directed self-interested concern with regard to that experience than it is like believing that you will have that experience. But first-personally anticipating an experience is not exactly the same as having future-directed self-interested concern with regard to that experience.

To see why I say that they are not exactly the same, suppose that you are about to have the experience of entering an even-numbered classroom (e.g. room 120). Add that having this particular experience will be neither good for you nor bad for you. Future-directed self-interested concern involves experiences that will be good or bad for you. So you lack future-directed self-interested concern with regard to entering an even-numbered classroom. But add that you first-personally anticipate entering an even-numbered classroom. So first-personally anticipating an experience is not exactly the same as having future-directed self-interested concern with regard to that experience.

I said that future-directed self-interested concern involves experiences that will be good or bad for you. A better account of future-directed self-interested concern would consider not only a future

[1] The way that Leibniz presents the king of China thought experiment makes it clear that he rejects a memory criterion of personal identity over time. Moreover, Leibniz says: 'So it is not memory that makes the same man' and ' . . . there is a perfect bond between the future and the past, which is what creates the identity of the individual. Memory is not necessary for this, however . . . ' (*New Essays on Human Understanding*, Bk II, 115 [1996]).

Here is one argument that Leibniz gives against the memory criterion. There is a possible situation in which you have a psychological duplicate on another planet. The memory criterion implies that, in this situation, that extraterrestrial duplicate is identical with your earlier self. You are identical with your earlier self. So—given the memory criterion—you are now thereby identical with your extraterrestrial duplicate. But Leibniz says: 'that would be a manifest absurdity' (*New Essays on Human Understanding*, Bk II, 245 [1996]).

experience's being good (or bad) for you, but also your believing that that experience will be good (or bad) for you; its being the case that, for all you know, that experience might (or might not) be good (or bad) for you; its being the case that you prefer to have (or prefer not to have) that experience, regardless of whether it will be good (or bad) for you; and so on. This better account would still imply the above point that first-personal anticipation is not exactly the same as future-directed self-interested concern. And that is the main point here. That point is consistent with my keeping things simple by oversimplifying. So I shall keep saying that future-directed self-interested concern involves experiences that will be good (or bad) for you.

First-personal anticipation is not exactly the same as future-directed self-interested concern. But they are closely related. Having future-directed self-interested concern with regard to an experience implies first-personally anticipating that experience. And first personally anticipating a good (or bad) experience implies having future-directed self-interested concern with regard to that experience. (Perhaps first-personal anticipation is a component of future-directed self-interested concern.)

Many contemporary philosophers are interested in appropriate first-personal anticipation or in appropriate future-directed self-interested concern. For example:

David Velleman: 'What we most want to know about our survival, I believe, is how much of the future we are in a position to anticipate experiencing.' (1996, 67)

Marya Schechtman: 'survival, moral responsibility, self-interested concern, and compensation . . . are indeed linked to facts about personal identity, but identity in the sense of the characterization question, not the reidentification question.' (1996, 2)

Eric Olson: 'Ultimately it is for ethicists to tell us when prudential concern is rational, when someone can be held accountable for which past actions, and who deserves to be treated as whom.' (1997, 70)

Jennifer Whiting: 'My general view is that the numerical identity of our present and future selves . . . is irrelevant to the justification of concern for our future selves.' (1986, 548)

Whiting elsewhere talks about 'the rationality of concern for oneself' (1991, 3) and 'the rationality of prudence' (1991, 3). And Jeff

McMahan (2002, 77–82) asks when 'egoistic concern' with regard to a future experience is 'rational'.

These philosophers also have views about what would render first-personal anticipation or future-directed self-interested concern 'rational' or 'justified' and views about when you are 'in a position' to first-personally anticipate, or have self-interested concern with regard to, a future experience. So they have views about the answer to this question:

> *The Wordier Why Question*: What way of being related to a (conscious) person at a future time explains why it is appropriate for you to both (i) first-personally anticipate the experiences that that person will have at that time and also (ii) have future-directed self-interested concern with regard to the good (or bad) experiences that that person will have at that time?

As we have seen, Leibniz's answer is that that person must not only be you, but must also, at that future time, have memories of your current life. I myself answer in terms of numerical identity (see Ch. 2, §I). McMahan (2002, 73 and 321) answers in terms of a combination of physical and functional continuity and psychological continuity. Velleman, Schechtman, and Whiting answer in terms of some sort of psychological connectedness or other (see Chs 4 and 5).[2] But none of us answers this question in terms of evidence. And while Olson (1997, 70) does not answer this question at all, he does say that the answer should come from ethicists, not epistemologists.

So let us assume that the Wordier Why Question should not be answered in terms of evidence. Then we should conclude that the relevant sort of appropriateness—the sort of appropriateness asked about in the Wordier Why Question—is not evidential.

We can reinforce the point that the relevant sort of appropriateness is not evidential by considering a case of first-personal anticipation and future-directed self-interested concern in which an evidential norm is violated. That case starts with Jones, who has won the fabled Nobel

[2] Psychological connectedness is constituted by, among other things, remembering an experience or being psychologically alike in some way. Psychological continuity is constituted by a chain of overlapping instances of psychological connectedness.

Prize in Philosophy. She knows that she has won, and she knows that she is slated to receive the Prize next week at a ceremony in Sweden. Jones is right now first-personally anticipating—and extending future-directed self-interested concern to—receiving the Prize next week.

Brown has not won, and is not slated to receive, the Nobel Prize in philosophy. Nor has Brown encountered (misleading) evidence that he has won the Prize. Brown has no reason at all to believe that he will receive the Prize. Brown is not even a philosopher! But Brown, no less than Jones, is right now first-personally anticipating—and extending future-directed self-interested concern to—receiving the Prize next week.[3]

Let us agree that it is not 'evidentially appropriate' for the non-philosopher Brown to first-personally anticipate, or have future-directed self-interested concern with regard to, receiving the Nobel Prize in Philosophy next week. For this to become evidentially appropriate, Brown would need some evidence. But evidence of what?

I think that Brown needs evidence that (someone identical with) Brown will receive the Prize. But Leibniz would disagree with me. Leibniz might say, instead, that Brown needs evidence that Brown will have memories of his current life when he receives the Prize. Velleman, Schechtman, and Whiting would also disagree with me. They might say, instead, that Brown needs evidence that he will be relevantly psychologically connected to the Prize recipient.

But there is something we can all agree on. We can all agree that Brown needs evidence that he will be related to the Prize recipient in whatever way makes it appropriate—in the non-evidential way at issue in this section—for him to first-personally anticipate, and have future-directed self-interested concern with regard to, receiving the Prize. So the evidential sort of appropriateness pertaining to first-personal anticipation and future-directed self-interested concern must be

[3] Brown can first-personally anticipate receiving the Prize even though he will not receive the Prize. For it is not a conceptual truth that one first-personally anticipates experience E only if one will have experience E. To see this, consider that Jones's thus anticipating does not guarantee that she will not die tomorrow in a tragic accident, and so fail to receive the Prize next week. A parallel point holds for having future-directed self-interested concern.

understood (in part) in terms of the sort of appropriateness on which this section has focused.

The same goes for the moral sort of appropriateness. Pretend that the Moral Oracle tells you that only concern for others is morally permitted, never concern for yourself. Suppose this implies that it is immoral to have self-interested concern with regard to your future experiences. (Or suppose that this implies that it is immoral to have self-interested concern with regard to the experiences that will be had by someone with whom you will be relevantly psychologically connected.) Then the following is true, and so not contradictory: it is not morally appropriate for you to have self-interested concern with regard to those future experiences with regard to which it is appropriate (in the way at issue in this section) for you to have self-interested concern.

Conversely, suppose that it is morally appropriate to have self-interested concern with regard to your future experiences. (Or suppose that it is morally appropriate to have self-interested concern with regard to the experiences that will be had by someone with whom you will be relevantly psychologically connected.) Then the following is not only true, but also non-trivial: it is morally appropriate for you to have self-interested concern with regard to those future experiences with regard to which it is appropriate (in the way at issue in this section) for you to have self-interested concern. So the moral sort of appropriateness pertaining to future-directed self-interested concern must be understood (in part) in terms of the sort of appropriateness on which this section has focused. And, for the rest of the book, I shall use the word 'appropriate' to mean appropriate in just this way.

Again, I shall use the word 'appropriate' for this non-evidential, non-moral norm. As we saw above—and shall see throughout this book—others often use the words 'justified' and 'rational' for this norm. But I myself think that 'justified' can misleadingly suggest an evidential norm, or at least an epistemic norm. I also think that 'rational' can be misleading. To see why, suppose that Brown knows that the Death Star will destroy the earth (and so kill us all, Brown included) unless Brown first-personally anticipates receiving the Nobel Prize in philosophy. Then it is rational—in some sense of 'rational'—for Brown to first-personally anticipate receiving the Prize. This is rational even though Brown is not himself a philosopher, will not receive the Prize, will not be psychologically connected with the person who will receive the Prize,

and so on. That is, Brown's first-personally anticipating receiving the Prize is rational even though it is not appropriate for Brown to first-personally anticipate receiving the Prize.

I have emphasized that the relevant sort of appropriateness is not an evidential (or other epistemic) norm or a moral norm or a matter of the sort of rationality just considered. Focusing on the fact that the relevant sort of appropriateness is not this or that familiar norm might make you suspect that it is no norm at all. So let me close this section by reminding you what role this norm plays, and so why there must be a norm that plays this role.

Leibniz would say that if you will lose all memories of your previous life and then have some regal experiences, it is not appropriate for you now to first-personally anticipate, or have future-directed self-interested concern with regard to, those experiences. Imagine someone opposing Leibniz by saying that if you will have regal experiences, then it is appropriate for you now to first-personally anticipate, and have future-directed self-interested concern with regard to, those experiences; Leibniz's opponent here adds explicitly that this is appropriate regardless of what you will remember when having those experiences. (I myself oppose Leibniz in just this way (see Ch. 2, §I).)

This disagreement between Leibniz and his opponent is intelligible. And to recognize that this disagreement is intelligible is to locate the relevant norm. For Leibniz and his opponent are not disagreeing about whether it is psychologically possible to first-personally anticipate, or have future-directed self-interested concern with regard to, an experience you will have after forgetting all that you have been. They are instead disagreeing about whether this is what you should do. But not morally. Or epistemically. Or Death-Star-rationally. Instead, they are disagreeing about whether your first-personally anticipating, and having future-directed self-interested concern with regard to, your post-amnesia regal experiences satisfies the norm that throughout this book will be indicated with the word 'appropriate'.

II. My Answer to the What Question

Let us ask:

> *The What Question*: What is it for a person at a future time to have (at that time) what matters in survival for you?

My answer: first, its being appropriate for you to first-personally anticipate the experiences that that person will have at that future time; and, second, if that person will have good (or bad) experiences at that time, its being appropriate for you to have future-directed self-interested concern with regard to that person's good (or bad) experiences at that time.

As we shall see below, my answer to the What Question is not particularly controversial. And I think we can already see that my answer is quite natural. For suppose that you will forget all that you have been and then become the king of China. Leibniz would say that it is not appropriate for you to first-personally anticipate, or have future-directed self-interested concern with regard to, the experiences that that king will have. It would be quite natural to describe Leibniz's diagnosis thus: even though that king will be you, that king will not have what matters in survival for you.

It makes sense to say that your giving money to charity is the right thing to do. And this still makes sense even if you do not give money to charity. Similarly, it makes sense to say that your first-personally anticipating an experience is appropriate, and this still makes sense even if you do not first-personally anticipate that experience. The same goes for having future-directed self-interested concern. So my answer to the What Question is consistent with a person's having, at a future time, what now matters in survival for you even if you do not now first-personally anticipate, or have future-directed self-interested concern with regard to, the experiences that that person will have at that time.

Let me illustrate this with a modified version of the story of Jones and the Prize. Jones has won the Prize. And Jones will receive the Prize in a few weeks. And she will suffer no metaphysical misadventure (no amnesia, no fission . . .) between now and receiving the Prize. But Jones does not yet know that she has won. (The phone is about to ring.) So Jones does not yet have future-directed self-interested concern with regard to receiving the Prize. Nevertheless, it is now appropriate for Jones to have future-directed self-interested concern with regard to receiving the Prize.[4] After all,

[4] Of course, it is not now evidentially or otherwise epistemically appropriate for Jones to have future-directed self-interested concern with regard to receiving the Prize. (The phone has not yet rung.) But the sense of 'appropriate' relevant to answering the What Question is not epistemic. See §I.

receiving the Prize next week will be a good thing for Jones, that is, for the Jones of right now.

Again, receiving the Prize next week will be a good thing for Jones, that is, for the Jones of right now. This suggests an alternative answer to:

> *The What Question*: What is it for a person at a future time to have (at that time) what matters in survival for you?

The alternative answer: that person's experiences at that future time will be good (or bad) for you, that is, for the you of right now.

The alternative answer to the What Question is closely related to my answer. For suppose that a person's future experiences will be good (or bad) for the you of right now. Then it is now appropriate for you to have future-directed self-interested concern with regard to those experiences. So it is now appropriate for you to first-personally anticipate those experiences. All this brings us back to my answer to the What Question.

Conversely, my answer to the What Question can lead us right to the alternative answer, at least when good (or bad) experiences are involved. For suppose that it is now appropriate for you to first-personally anticipate the experiences that a person will have at a future time. Add that those experiences will be good (or bad). Then it is now appropriate for you to have future-directed self-interested concern with regard to those experiences. So it must be that those experiences will be good (or bad) for you, that is, for the you of right now. This is the alternative answer to the What Question.

We have just seen that my answer to the What Question is closely related to the alternative answer. As a result of how they are related, you can agree with the most important arguments and conclusions in this book even if you accept the alternative answer to the What Question in place of my answer. This should be clear in what follows, especially because I shall often return to the point that a future (conscious) person's good (or bad) experiences will be good (or bad) for you just in case that person will have what matters in survival for you. Nevertheless, there are three reasons that I prefer my answer to the alternative answer.

The first reason is that I think that my answer is more standard than the alternative answer. For many philosophers take appropriate

first-personal anticipation or appropriate future-directed self-interested concern to be constitutive of what matters in survival. For example, McMahan (2002, 43) describes Derek Parfit's (1984) treatment of what matters in survival as a treatment of 'rational egoistic concern'. Similarly, Sydney Shoemaker (1985, 444) takes Parfit's views about what matters in survival to be views about 'what makes it rational (to the extent that it is) for me to have a special concern for my well-being'.

The second reason begins by recalling that Leibniz, Whiting, Schechtman, Olson, Velleman, and McMahan focus on appropriate first-personal anticipation or appropriate future-directed self-interested concern (§I). So do other philosophers, as we shall see in later chapters. I want it to be easy to see how what I say about what matters in survival interacts with what these philosophers say about appropriate first-personal anticipation and about appropriate future-directed self-interested concern. My answer to the What Question makes this easier to see than does the alternative answer to the What Question.

There is a third reason. Suppose nothing good or bad happens to a person at a future time. (Perhaps that person is simply entering an even-numbered room at that time.) Then that person's experiences at that time will not be good (or bad) for anyone. So they will not be good (or bad) for you, not even for the you of right now. But we ought to be able to accommodate the possibility that a person will have what matters in survival for you at a future time even if that person will not be having a good (or bad) experience at that time. My answer to the What Question accommodates this possibility, since it can be appropriate for you to first-personally anticipate a person's experiences at a future time even if those experiences are not good (or bad). The alternative answer to the What Question cannot accommodate this possibility, since that answer takes what matters in survival to be constituted by a person's experiences at a future time being good (or bad) for you, that is, for the you of right now.

So I have three reasons for preferring my answer to the What Question to the alternative answer. My third reason turns on the possibility of surviving as a person at a future time even if nothing good (or bad) will happen to that person at that time. This possibility shows not only that we should reject the above alternative answer to the What Question, but also that we should not answer that question exclusively in terms of appropriate future-directed self-interested concern.

We should not answer the What Question in terms of appropriate future-directed self-interested concern alone. But you might answer the What Question in terms of appropriate first-personal anticipation alone. Then your answer is in some sense equivalent to my answer. This is because its being appropriate to first-personally anticipate a good (or bad) experience implies its being appropriate to have future-directed self-interested concern with regard to that experience, and vice versa. Moreover, your answer has the advantage of comparative concision over my answer.

Nevertheless, I prefer my answer. Recall that I want it to be easy to see how what I say about what matters in survival interacts with what others say about appropriate future-directed self-interested concern. Explicitly including such concern in my answer to the What Question makes this easy to see. And making this easy to see is more important to me than answering the What Question as concisely as possible.

In remarks quoted in Section I, Olson (1997, 70) mentions being 'accountable for . . . past actions' and Schechtman (1996, 2) mentions 'moral responsibility' and 'compensation'. Let me explain why my answer to the What Question does not include moral responsibility or just compensation or accountability for past actions.

Suppose that moral responsibility (and so on) can pull apart from appropriate first-personal anticipation and appropriate future-directed self-interested concern (cf. Olson, 1997, 68; Shoemaker, 2016). Then I would say that appropriate first-personal anticipation and appropriate future-directed self-interested concern constitute what matters in survival even in the absence of future moral responsibility (and so on) for present acts.

Alternatively, suppose that moral responsibility (and so on) cannot pull apart from appropriate first-personal anticipation and appropriate future-directed self-interested concern. Then I would say that what explains moral responsibility (and so on) is the having of what matters in survival. For example, I would say that there is a reason that you are morally responsible for a past action, and that reason is that the person who thus acted has survived as you; that is, that you now have what mattered in survival for that person at the time that that person acted.

I have been motivating my answer to:

> *The What Question*: What is it for a person at a future time to have (at that time) what matters in survival for you?

But when it comes to the main conclusions of this book, my answer to the What Question is less important than my answer to:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

Indeed, when it comes to the main conclusions of this book, I could have skipped both the What Question and also the Why Question, and instead asked and answered:

> *The Wordier Why Question*: What way of being related to a (conscious) person at a future time explains why it is appropriate for you to both (i) first-personally anticipate the experiences that that person will have at that time and also (ii) have future-directed self-interested concern with regard to the good (or bad) experiences that that person will have at that time?

But suppose that I had skipped both the What Question and the Why Question and instead asked only the Wordier Why Question. Then I would not have motivated my answer to the What Question. Then it might not be clear that an answer to the Wordier Why Question is relevant to the value of personal identity, or to what matters in survival. For my motivation of my answer to the What Question was meant, in part, to convince you that an answer to the Why Question—and so to the Wordier Why Question—is relevant to the value of personal identity and to what matters in survival. And the philosophers discussed above should definitely be convinced. For my answer to the What Question is inspired by their own views about the value of personal identity and what matters in survival.

Moreover, I could not have explicitly distinguished the What Question from the Why Question if I had skipped those two questions and instead asked only the Wordier Why Question. But I think it is important to explicitly distinguish the What Question from the Why Question. To begin to see why I think this is important, recall a point made in the Introduction: someone might conflate the What Question and the Why Question.

Let me now add that those questions seem to have actually been conflated. For those questions sometimes seem to be given the same answer. Giving the same answer to both of those questions is

tantamount to saying that a future person's having what matters in survival for you is explained by that person's having what matters in survival for you. But I doubt that anyone really wants to say that. So I suspect that anyone who gives the same answer to both the What Question and the Why Question has conflated those questions.[5]

The What Question and the Why Question might be—and I think sometimes are—conflated. And this conflation is not harmless. For this conflation can easily lead to bad reasoning about what matters in survival. The following is an example of such bad reasoning, an example that is of particular importance in light of Chapter 2: numerical identity is not what matters in survival; therefore, it is false that your being numerically identical with a person at a future time explains why that person will have, at that time, what matters in survival for you.

III. Consciousness and Survival

Thomas Nagel (1970, 74) says: 'almost everyone would be indifferent (other things equal) between immediate death and immediate coma followed by death twenty years later without reawakening'. But no one would be indifferent between death and permanent coma if the comatose were known to be conscious. So at least one reason for the indifference Nagel points to must be the belief that the comatose are not conscious.[6] I take this indifference to go along with the idea that no permanently comatose person will have what matters in survival for you. So I conclude that at least one reason that no permanently comatose person will have what matters in survival for you is that the comatose are not conscious.

[5] For instance, Derek Parfit (1971, 20; 1984, 262) sometimes seems to answer both (what I call) the What Question and (what I call) the Why Question in terms of psychological connectedness and/or continuity. This is discussed in Ch. 3 (§II). As we shall see below (§IV), Parfit sometimes seems to answer the What Question in a couple of other ways as well.

[6] I share this belief. But there is evidence that those in a persistent vegetative state have conscious experiences (see Cryanoski (2012)).

Consider two controversial, but defensible, theses: you are identical with a human organism; and after that human organism dies, it will continue to exist awhile as a corpse.[7] These theses jointly imply that you will one day be a corpse. But I say that no corpse will have what matters in survival for you, or for anyone. On the other hand, if corpses were conscious, then I think that a future corpse could have what matters in survival for you. So I conclude that at least one reason that no corpse will have what matters in survival for you is that corpses are not conscious.

No permanently comatose person will have what matters in survival for you because the comatose are not conscious. No corpse will have what matters in survival for you because corpses are not conscious. So I conclude that an entity will have, at a future time, what matters in survival for you only if that entity will be conscious at that future time.

This conclusion seems right to me. And it will seem right to others as well. For example, consider this passage from Velleman, part of which was quoted above (§I):

> What we most want to know about our survival, I believe, is how much of the future we are in a position to anticipate experiencing. We peer up the stream of consciousness, so to speak, and wonder how far up there is still a stream to see. (1996, 67–8)

Suppose that a person who is in dreamless sleep at a future time is not conscious at that time.[8] Then I say that no person in dreamless sleep at a future time has, at that time, what matters in survival for you, or for anyone else. This is less striking than it might seem. This is because usually a person in dreamless sleep will eventually wake up (or dream) and so will eventually have what matters in survival for someone. On the other hand, if such a person will never wake up (or dream), it seems right that that person will never have what matters in survival for anyone. For when it comes to what matters in survival, entering permanent and unconscious sleep seems to be on a par with entering permanent and unconscious coma.

[7] I reject the second thesis, denying that there are any corpses (as opposed to *x*s arranged corpsewise; see Merricks, 2001a, 53).

[8] This supposition is controversial. See Windt, Nielsen, and Thompson (2016).

Return to:

> *The What Question*: What is it for a person at a future time to have (at that time) what matters in survival for you?

And recall my answer: its being appropriate for you to first-personally anticipate the experiences that that person will have at that future time; and if that person will have good (or bad) experiences at that future time, its being appropriate for you to have future-directed self-interested concern with regard to those experiences.

I intend the experiences invoked in my answer to the What Question to be conscious experiences. So take those experiences to be conscious experiences. Then my answer has the result that a person at a future time will have, at that time, what matters in survival for you only if that person will be conscious at that time. As I have argued in this section, I think that this is the right result.

Recall:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

I shall argue in Chapter 2 that being numerically identical with is a good answer to the Why Question. And I shall sometimes summarize the view that numerical identity is a good answer to the Why Question with the following slogan: identity delivers survival. A less pithy but more accurate slogan would be: identity with a conscious person delivers survival.

Pretend that something bad will happen to you while you are permanently unconscious and comatose. Then I think that it is appropriate for you to have self-interested concern with regard to that bad happening. I say that what makes this appropriate is your being numerically identical with that comatose person. So I am not claiming that a future person's being conscious is necessary for it to be appropriate for you to have self-interested concern with regard to what that person will go through. Instead, I am claiming that a future person's being conscious is necessary for that person's having what matters in survival for you.

If something bad will happen to you while you are permanently unconscious and comatose, this will be bad for you, that is, for the you

of right now. This will be bad for you even though—because you will not then be conscious—no one will have what matters in survival for you when this happens. So it is false that something that involves a person at a future time will be bad (or good) for you only if that person will have, at that time, what matters in survival for you. But this is false only because that person might, at that time, fail to be conscious.

Suppose that something good (or bad) will happen to a conscious person. Then that something will be good (or bad) for you—that is, for the you of right now—only if that person will have what matters in survival for you. In what follows, I shall often say of an example involving a person at a future time that what happens in this example will be good (or bad) for you only if that person will have, at that time, what matters in survival for you (see, esp., Ch. 4, §§IV–V; Ch. 5, §§IV–V). That is fine. For all these examples involve a person who is conscious at that time.

You might deny that a person at a future time will have, at that time, what matters in survival for you only if that person will be conscious at that time. I disagree. But our disagreement on this single point really is just disagreement on this single point. For even given this disagreement, you can still endorse my answers to the What Question and the Why Question, as well as my arguments for those answers. But you should take the experiences invoked in my answer to the What Question to be all that one goes through, which includes more than having conscious experiences. And you may drop the parenthetical 'conscious' from the Why Question. So you can take the slogan 'identity delivers survival' to be as accurate as it is pithy.

IV. What Matters to You with Regard to the Future

As we saw above (§§I–II), Leibniz thinks that a future person will have what matters in survival for you just in case that person will not only be you but will also remember 'what you have been'. Leibniz adds:

> For it is memory or the knowledge of this self that renders it capable of punishment or reward. ('Discourse on Metaphysics', §34 [1989, 66])

So Leibniz would say that a person who will have, at a future time, what matters in survival for you can be punished at that time. But let us

add that avoiding punishment at a future time is one of the things that matters to you with regard to that time.

Suppose that it is appropriate for you both to first-personally anticipate suffering at a future time, and also to have self-interested concern with regard to that suffering at that time. Then a suffering person will have, at that time, what matters in survival for you (§II). But let us add that avoiding suffering at a future time is another thing that matters to you with regard to that time.

These remarks about future punishment and future suffering illustrate the following point: a person at a future time's having, at that time, what matters in survival for you is not the same thing as that person's having, at that time, all that matters to you with regard to that future time. This point should be obvious. That is, it should be obvious that surviving is not the same thing as getting all that you want.

Surviving is not even the same thing as getting part of what you want. That is, a person at a future time having what matters in survival for you is not even the same thing as that person's having, at that time, part of what matters to you with regard to that time. For suppose that you have grown tired of life and you want it all to end. That is, suppose that you do not want there to be, at any future time, a person who will have, at that time, what now matters in survival for you. This is depressing. But it is not contradictory.

You probably do want to survive. That is, part of what probably matters to you with regard to a future time is that there will be someone who will have, at that time, what matters in survival for you. But this is a substantive fact about you. This is not a trivial result of the nature of what matters in survival. For, again, there is nothing contradictory about your not wanting anyone to have what matters in survival for you at a future time.

Suppose that it matters to you that the person who will have what matters in survival for you at a future time will have friends at that time. Or suppose that it matters to you that the person who will have what matters in survival for you at a future time will not be in grinding poverty at that time. These are claims about what matters to you with regard to that future time. But these claims are not equivalent to the claim that a person will have what matters in survival for you at that future time. Rather, these claims are understood partly in terms of a person's having what matters in survival for you at that time. Again, a

person at a future time having, at that time, what matters in survival for you is not the same thing as that person's having, at that future time, what matters to you with regard to that future time.

Derek Parfit makes many claims that he says are about what matters in survival. And at least some of those claims really are about (what I am calling) what matters in survival. For example:

> An emotion or attitude can be criticized for resting on a false belief or for being inconsistent. A man who regards [double brain hemisphere transplant] as death must, I suggest, be open to one of these criticisms. (Parfit, 1971, 9)

I think that Parfit is here claiming that if a man is going to undergo double brain hemisphere transplant, both of the resulting persons will have what matters in survival for that man (see Ch. 3, §II).

Moreover, Parfit (1984, 263) claims that if a person at a future time is the same person as you, then that person will have, at that future time, what matters in survival for you (see Ch. 3, §II). That is, Parfit claims that personal identity is sufficient (but not necessary) for what matters in survival. This too seems to be a claim about what matters in survival for you, as opposed to a claim about what matters to you with regard to a future time. For I do not think that Parfit is claiming that your being the same person as a person at a future time is sufficient for what matters to you with regard to that time. For example, it would be silly to say that being the same person as a person at a future time is sufficient for your having friends at that time.

But some of Parfit's claims that he says are about 'what matters in survival' are, instead, claims about what matters with regard to a future time. For example, Parfit claims that not having a doppelgänger compete for the affection of one's beloved at a future time 'matters in survival' (Parfit, 1984, 264). But this is surely a claim about what matters to one with regard to that future time.

And consider the following passage from Parfit, which conflates what matters in survival for him and what matters to him with regard to a future time:

> Just as division shows that what matters in survival need not take a one-one form, fusion shows that it can have degrees . . . The value to me of my relation to a resulting person depends both (1) on my degree of [psychological] connectedness to this person, and (2) on the value, in my view, of this person's physical and psychological features. Suppose that hypnosis causes me to lose

five unwanted features: my untidiness, laziness, fear of flying, nicotine addiction, and all my memories of my wretched life. There is here much less than full psychological connectedness, but this is more than outweighed by the removal of bad features. (1984, 298–9)[9]

A person at a future time's having, at that time, what matters in survival for you has been conflated with a person at a future time's having, at that time, what matters to you with regard to that time. This conflation is not harmless. For this conflation can lead to bad reasoning about how to answer:

> *The What Question*: What is it for a person at a future time to have (at that time) what matters in survival for you?

To begin to see why I say this, consider the following remarks from David DeGrazia:

> So far, our answer to the question of what . . . matters in survival has considered only *experience*. But to stress experience is to stress only a relatively passive side of human persons: what we take in through the senses and process with our minds. Of course, we humans are also agents—beings who act, sometimes spontaneously, sometimes after deliberation and planning. Agency seems no less central to what we are . . . , and what we care about, than experience is. (2005, 79)

Pretend that you have conflated what matters in survival and what matters with regard to the future. Then you might take DeGrazia's remarks to constitute an objection to my answer to the What Question, which is in terms of experiences (that are appropriately first-personally anticipated, etc.). But taking DeGrazia's remarks in this way would be a mistake. For it is a mistake to conflate what matters in survival and what matters with regard to the future.

Look at it this way. An exhaustive list of what matters to you with regard to a future time would include more than your having

[9] One result of this conflation's occurring in Parfit's work is that some objections to what Parfit says about 'what matters in survival' have nothing to do with what matters in survival. They are instead objections to Parfit's claims about what does (or should) matter to one with regard to the future (see, e.g. Wolf, 1986, 714–15).

experiences at that time, and even more than your having experiences that you can now appropriately first-personally anticipate (etc.). It would include your exercising agency. But it would also include your having friends, not being in grinding poverty, doing meaningful work, and much more. None of this suggests that an answer to the What Question should mention friends or money or work. So none of this suggests that an answer to the What Question should mention agency.[10]

I have used 'survive' above as shorthand for there being a person at a future time who will have, at that time, what matters in survival. I shall continue to do this—but with more frequency—for the rest of the book.

One reason for using 'survive' in this way is that it should help us to avoid the conflation identified in this section. For example, this false and conflating sentence might appear, at first glance, to be true: 'a person's having what matters in survival for you at a future time just is that person's having what matters to you with regard to that future time'. On the other hand, I do not think that this false and conflating sentence will appear, even at first glance, to be true: 'your surviving as a person at a future time just is that person's having at that time what matters to you with regard to that future time'.

Another reason for using 'survive' in this way is concision. For example, this allows us to replace 'there will be a person who will have what matters in survival for you' with 'you will survive'. And 'you will survive a change' can replace 'you will undergo a change and, after that change, there will be a person who will have what matters in survival for you'.

But I admit that there is a downside to using 'survive' as shorthand for there being a person at a future time who will have, at that time, what matters in survival. This downside is the danger of a new conflation, conflating what matters in survival and persistence. This

[10] The What Question should not be answered in terms of agency. But this is consistent with agency's having a special role to play in what matters in survival, a role that is not played by, for example, having friends. For this is consistent with answering the Why Question in terms of agency. Chapter 6 considers an answer to the Why Question in terms of agency.

is a danger because some philosophers use 'survive' to mean persist (see, e.g. DeGrazia, 2005, 79; Kagan, 2012, 2). So beware!

V. Conclusion

This chapter defended my answer to:

> *The What Question*: What is it for a person at a future time to have (at that time) what matters in survival for you?

My answer: its being appropriate for you to first-personally anticipate the experiences that that person will have at that future time; and if that person will have good (or bad) experiences at that future time, its being appropriate for you to have future-directed self-interested concern with regard to those experiences.

Obviously, my answer to the What Question is not that you are numerically identical with that person at that future time. Moreover, I never defended my answer to the What Question with any claims about the metaphysics of persistence. This is because the metaphysics of persistence is irrelevant to answering the What Question.

The next chapter will defend my answer to:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

If I had conflated the What Question and the Why Question, I might have taken my answer to the What Question to be the answer to the Why Question. I would have then concluded that the answer to the Why Question is not that you are numerically identical with that person at that future time. And I would also have concluded that the metaphysics of persistence is irrelevant to answering the Why Question. But those conclusions would have been based on a mistaken conflation (§II).

2

On the Sufficiency of Personal Identity

Let us ask:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

My answer is numerical identity. That is, I say that your being numerically identical with a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you.

This answer is controversial. For this answer implies the claim that your being numerically identical with a person at a future time explains why it is appropriate for you to first-personally anticipate, and have future-directed self-interested concern with regard to, the experiences that that person will have at that time. And many philosophers deny this claim. For example:

G. W. Leibniz: 'Thus the immortality required in morality and religion does not consist merely in this perpetual subsistence common to all substances, for without the memory of what one has been, there would be nothing desirable about it.' ('Discourse on Metaphysics', §34 [1989, 66])

Jennifer Whiting: 'My general view is that the numerical identity of our present and future selves . . . is irrelevant to the justification of concern for our future selves.' (1986, 548)

Marya Schechtman: 'Most modern personal identity theorists, I charge, conflate two significantly different questions, which I call the reidentification

Self and Identity. Trenton Merricks, Oxford University Press. © Trenton Merricks 2022.
DOI: 10.1093/oso/9780192843432.003.0003

question and the characterization question. The former is the question of what makes a person at time t_2 the same person as a person at time t_1; the latter the question of which beliefs, values, desires, and other psychological features make someone the person she is . . . . There is a strong pre-philosophical sense that facts about personal identity underlie facts about four basic features of personal existence: survival, moral responsibility, self-interested concern, and compensation . . . I contend that the four features are indeed linked to facts about personal identity, but identity in the sense of the characterization question, not the reidentification question.' (1996, 1–2)

David Velleman: 'What we most want to know about our survival, I believe, is how much of the future we are in a position to anticipate experiencing . . . what I want to know is a matter of perspective rather than metaphysics. My question is *not* how long there will be an individual identical with [David Velleman].' (1996, 67–8)

My answer to the Why Question is controversial. So is my way of defending that answer. For my defense of that answer turns on the metaphysics of persons. But many philosophers deny that the metaphysics of persons is relevant to what matters in survival, and so should deny that it is relevant to answering the Why Question.

Here is Velleman again:

The appeal of [the topic of personal identity] depends largely on its promise to address our concern about what we can look forward to, or what we can anticipate first-personally. If the mode of anticipation that arouses our concern is first-personal in the sense of being framed from the perspective of a future person, rather than in representing the future existence of the anticipator, then that concern should move us to study the psychology of perspectives rather than the metaphysics of persons. (1996, 41)

In a paper in which he motivates 'Minimalism', Mark Johnston says:

In the particular case of personal identity, Minimalism will imply that any metaphysical view of persons that we might have is not indispensable to the justification of our practice of making judgments about personal identity and organizing our practical concerns around these judgments. (1992, 590)

And Christine Korsgaard seems to be skeptical about the relevance of the metaphysics of persons to what matters in survival. Korsgaard says:

. . . the metaphysical facts do not obviously settle the question: I must still decide whether the consideration that some future person is 'me' has some

special normative force for me. It is practical reason that requires me to construct an identity for myself; whether metaphysics is to guide me in this or not is an open question. (1989, 112)

In this chapter, I defend the claim that numerical identity is a good answer to the Why Question. (But I never insist that numerical identity is the only good answer to the Why Question.) My defense of this claim relies on a view about the metaphysics of persons; in particular, it relies on the view that persons 'endure'. So this chapter defends a controversial claim in a controversial way.

I. My Answer to the Why Question

Suppose you are now experiencing pain. Then it is appropriate for you now to have self-interested concern with regard to that experience of pain. Moreover, your now experiencing pain explains why it is appropriate for you now to have self-interested concern with regard to that experience of pain. There might be other good explanations as well. But the point here is that your now experiencing pain is one good explanation.

Your now experiencing pain implies that, and explains why, it is now appropriate for you to have self-interested concern with regard to that experience of pain. Moreover, your now having any good or bad experience implies that, and explains why, it is now appropriate for you to have self-interested concern with regard to that experience. And it is not just you. So:

> (0) For all persons x, x's now having good (or bad) experience E both implies that it is appropriate for x now to have self-interested concern with regard to E and also explains why this is appropriate.

And (0) remains true even when put in other words:

> (1) For all persons x and all persons y, x's being numerically identical with y and y's now having good (or bad) experience E both implies that it is appropriate for x now to have self-interested concern with regard to E and also explains why this is appropriate.

So (1) is true. And I think that if (1) is true, then so are parallel claims about what was and what will be. That is, if (1) is true, then so is the more general thesis that—whether this occurs in the past, present, or

future—a person's having experience E implies that, and explains why, it is appropriate for that person to have self-interested concern with regard to that experience. In other words—and whenever this occurs—a person's being numerically identical with the entity that has experience E implies that, and explains why, it is appropriate for that person to have self-interested concern with regard to that experience. So, for example:

> (2) For all persons x and all persons y, x's being numerically identical with y and the fact that y will have good (or bad) experience E at a future time both implies that it will be appropriate at that future time for x to have self-interested concern with regard to E and also explains why this will be appropriate at that future time.

(2) is a claim about self-interested concern at a future time. But it is not a claim about future-directed self-interested concern. Do (1) and (2) (and the more general thesis) motivate any claim about future-directed self-interested concern? It depends.

That is, it depends on whether a single entity both has future-directed self-interested concern with regard to an experience E and also will have experience E. In other words, it depends on whether the entity that has future-directed self-interested concern with regard to E is numerically identical with the entity that will have experience E. With this in mind, I take (1) and (2) (and the more general thesis) to motivate:

> (3) For all persons x and all persons y, x's being numerically identical with y and the fact that y will have good (or bad) experience E both implies that it is appropriate for x now to have future-directed self-interested concern with regard to E and also explains why this is appropriate.

I find this motivation for (3) compelling. So I endorse (3).[1]

[1] Because our topic is survival, I am focusing on future-directed self-interested concern. But let 'past-directed self-interested concern' be self-interested concern with regard to what has already happened. Obviously, past-directed self-interested concern does not involve first-personal anticipation. For first-personal anticipation is a matter of looking ahead to future experiences. Past-directed self-interested concern involves, instead, looking back at past experiences. Past-directed

But (1) and (2) (and the more general thesis) do not motivate just any claim to the effect that personal identity explains the appropriateness of future-directed self-interested concern. To see why I say this, consider, for example:

> (4) For all x and all y, x's being a temporal part of the same person as is y (but being numerically distinct from y), y's being located at a later time than is x, and y's having good (or bad) experience E both implies that it is appropriate for x now to have future-directed self-interested concern with regard to E and also explains why this is appropriate.

Unlike (3), (4) does not make a claim about an entity that has future-directed self-interested concern with regard to E and is numerically identical with the entity that will have experience E. So I deny that (1) and (2) (and the more general thesis) motivate (4). But (4) is a claim to the effect that personal identity explains the appropriateness of future-directed self-interested concern. According to (4), the appropriateness of future-directed self-interested concern is explained by the relevant temporal parts being parts of one and the same person. (More about temporal parts in §II.)

If persons *endure*, then the following are true. First, the relation of 'identity over time' is numerical identity. Second, persons are the relata of identity over time in cases of personal identity over time. Third, persons have 'temporary' properties, that is, properties with regard to which persons change; in particular, persons have future-directed self-interested concern with regard to experiences, and also have experiences.

Suppose that you will endure until a future time (and remain a person). This implies that you are numerically identical with a person who will exist at that future time. Add that that person will have a good (or bad) experience at that future time. Invoke:

self-interested concern might even just be (something like) remembering having good (or bad) experiences, or remembering having experiences that you prefer that you had (or had not had), and so on (cf. Ch. 1, §I). And I do endorse: for all persons x and all persons y, x's being numerically identical with y and the fact that y had experience E both implies that it is appropriate for x now to remember having E and also explains why this is appropriate.

> (3) For all persons x and all persons y, x's being numerically identical with y and the fact that y will have good (or bad) experience E both implies that it is appropriate for x now to have future-directed self-interested concern with regard to E and also explains why this is appropriate.

Then conclude that it is now appropriate for you to have self-interested concern with regard to that experience at that future time.

Suppose that you learn that tomorrow you will suffer through a painful dental procedure. One natural way for you to react to this news is by extending your self-interested concern to that suffering tomorrow. In light of (3)—and if you endure—I conclude that your natural reaction is appropriate, and that its appropriateness is explained by your being (numerically identical with) the person who will suffer.

Or suppose that you will become a king, but only after you lose all memories of your current life.[2] I think that you endure. So I conclude that it is appropriate for you to have future-directed self-interested concern with regard to that king's experiences. So it is appropriate for you to first-personally anticipate that king's experiences (Ch. 1, §II). This is all explained—in light of (3)—by your being numerically identical with that future king.

More generally, I conclude that being numerically identical with is a good answer to:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

This more general conclusion presupposes my answer to:

> *The What Question*: What is it for a person at a future time to have (at that time) what matters in survival for you?

[2] Those who endorse a criterion of personal identity over time in terms of memory connectedness would deny that you could persist as someone who will have lost all memories of your current life. But endurantists in particular have a compelling reason to reject that criterion: numerical identity is transitive but being connected by memory is not.

That answer, recall, is in terms of appropriate first-personal anticipation and appropriate future-directed self-interested concern (Ch. 1, §II).

The main point I have just defended could be defended without presupposing my answer to the What Question, and even without mentioning the Why Question. That main point is that being numerically identical with is a good answer to:

> *The Wordier Why Question*: What way of being related to a (conscious) person at a future time explains why it is appropriate for you to both (i) first-personally anticipate the experiences that that person will have at that time and also (ii) have future-directed self-interested concern with regard to the good (or bad) experiences that that person will have at that time?

But for the reasons given in Chapter 1 (§II), I shall continue to focus on both the What Question and the Why Question, as opposed to only on the Wordier Why Question.[3]

I think that a natural answer to the Wordier Why Question is in terms of numerical identity. That is, it is natural—and initially plausible—to say that it is appropriate for you to first-personally anticipate, and have future-directed self-interested concern with regard to, certain future experiences because it is you yourself who will have those experiences. By the same token, my answer to the Why Question is natural and initially plausible. But, as we saw at the start of this chapter, my answer to the Why Question is also very controversial. So I have defended it.

II. More on the Metaphysics of Persistence

In Section I, I argued that being numerically identical with is a good answer to:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

[3] Others who would take being numerically identical with to be a good answer to the Wordier Why Question include Roderick Chisholm (1969, 138; 1976, 113) and Harold Langsam (2001).

The argument in Section I requires the following three claims to be true. First, the relation of 'identity over time' is numerical identity. Second, persons are the relata of identity over time in cases of personal identity over time. Third, persons have temporary properties (i.e. properties with regard to which persons change); in particular, persons have future-directed self-interested concern with regard to experiences, and also have experiences.

Endurantists accept all three claims. But—as I shall argue in this section—only endurantists should accept all three claims. So—as I shall argue in this section—only endurantists should endorse Section I's argument for numerical identity's being a good answer to the Why Question. Suppose this section's arguments fail. Suppose that not only should endurantists endorse Section I's argument, but others should endorse that argument as well. Then others should join endurantists in concluding that numerical identity is a good answer to the Why Question. That would be great! So I hope that this section's arguments fail. But, sadly, I think they succeed.

If persons *perdure*, then the following are true. First, persons are four-dimensional and have temporal parts. Second, the temporal parts of a person have that person's temporary properties (i.e. the properties with regard to which that person changes); in particular, a person's temporal parts have future-directed self-interested concern with regard to experiences, and also have experiences.[4]

Recall:

> (3) For all persons x and all persons y, x's being numerically identical with y and the fact that y will have good (or bad) experience E both implies that it is appropriate for x now to have future-directed self-interested concern with regard to E and also explains why this is appropriate.

Perdurantists say that a person's temporal parts have future-directed self-interested concern. So I think that perdurantists should deny that persons themselves have future-directed self-interested concern. (3) tells us what explains the appropriateness of a person's having

[4] Almost all self-described perdurantists would agree with what I have just said about perduring persons. But not Josh Parsons (2000).

future-directed self-interested concern. So I think that perdurantists should deny that (3) tells us what explains the appropriateness of any actual instances of future-directed self-interested concern. The argument in Section I relies on (3). So I think that perdurantists should deny that Section I shows that numerical identity is a good answer to the Why Question that applies to actual cases of survival.

Some perdurantists will object. These perdurantists claim that persons do have future-directed self-interested concern. In particular, these perdurantists claim that a perduring person has future-directed self-interested concern with regard to an experience in virtue of having a temporal part that has future-directed self-interested concern with regard to that experience. And these perdurantists will add that the numerically same person can have that experience in virtue of having another temporal part that has that experience.

Suppose these perdurantists endorse the following:

> (4D3) For all persons x and all persons y, x's being numerically identical with y and the fact that y has good (or bad) experience E (in virtue of having a later temporal part that has experience E) both implies that it is appropriate for x to have future-directed self-interested concern with regard to E (in virtue of having an earlier temporal part that has future-directed self-interested concern with regard to E) and also explains why this is appropriate.

These perdurantists might take (4D3) to imply:

> (3) For all persons x and all persons y, x's being numerically identical with y and the fact that y will have good (or bad) experience E both implies that it is appropriate for x now to have future-directed self-interested concern with regard to E and also explains why this is appropriate.

So these perdurantists might take (3) to tell us what explains the appropriateness of actual cases of future-directed self-interested concern. So these perdurantists might conclude, in light of Section I, that numerical identity is a good answer to the Why Question that applies to actual cases of survival.

Perdurantists can take (3) to be true—and to tell us what explains the appropriateness of actual cases of future-directed self-interested concern—only if they endorse (4D3). And they can endorse (4D3)

only if they claim that a person has future-directed self-interested concern in virtue of having a temporal part that has such concern, and a person has an experience in virtue of having a temporal part that has that experience. But no perdurantist should claim this. So no perdurantist should endorse (4D3). I say this for two reasons.

The first reason starts with this schema: for every person P and every property F, if P has a temporal part that is F, then P is F (in virtue of having a temporal part that is F). This schema is false. To see this, let F be *being instantaneous* (or *being a proper part of P* or *not being a person* or . . .). A false schema can have true instances. But I do not see a principled way to motivate the truth of the instances of this false schema that let F be *has future-directed self-interested concern with regard to experience E* or that let F be *has experience E*. So I do not think there is a principled way to motivate (4D3).

This first reason considered a particular false schema. That schema was not the following schema: for every person P and every property F, if P has a temporal part at t that is F, then P is F-at-t (in virtue of having a temporal part at t that is F). Suppose that a perduring person has both a temporal part at t that has future-directed self-interested concern with regard to an experience E and also a later temporal part at t* that has experience E. Then this latter schema implies that that perduring person has future-directed-self-interested-concern-with-regard-to-E-at-t and has E-at-t*. But this latter schema does not imply that that perduring person has future-directed self-interested concern with regard to E, or that that person has experience E. Indeed, this latter schema is perfectly consistent with the view that only temporal parts—but never persons—have experiences and have future-directed self-interested concern with regard to experiences. So this latter schema does not lead to (4D3).

There is a second reason that perdurantists should not claim that a person has future-directed self-interested concern in virtue of having a temporal part that has such concern, or that a person has an experience in virtue of having a temporal part that has that experience. So there is a second reason that perdurantists should not endorse (4D3). This reason begins by observing that perdurantists typically take 'P does not have experience E at time t and P does have experience E at time t*' to amount to the claim that P has a temporal part located at time t that does not have E and P has a numerically distinct temporal part located

at time t* that does have E. This claim is manifestly non-contradictory. And it illustrates the way in which perdurance is widely taken to render change non-contradictory (see Haslanger, 2003, 331–4).[5]

But suppose that 'P does not have experience E at t and P has experience E at t*' amounts to the claim that P does not have experience E (in virtue of having a temporal part located at t that does not have E) and P does have experience E (in virtue of having a numerically distinct temporal part located at t* that does have E). This claim implies that P does not have E and that P does have E. This implication is not manifestly non-contradictory.

(4D3) makes a claim about having an experience. But (4D3) does not explicitly mention not having an experience. So some perdurantists might think that they can endorse (4D3) but deny that P's having a temporal part at t that does not have E implies that P does not have E. A quick—and decisive—reply to these perdurantists is to exchange the above example for one in which P's temporal part at t has experience F, the having of which is not consistent with having E. For instance, let E be having a pain-free experience, and F be being in excruciating pain. (A less quick reply would start with my above worry about a principled way to motivate (4D3).) But in what follows, I shall stick with the above example.

Perdurantists who endorse (4D3) cannot take 'P does not have experience E at time t and P does have experience E at time t*' to amount to the claim that P has a temporal part located at time t that does not have E and P has a numerically distinct temporal part located at time t* that does have E. More generally, perdurantists who endorse (4D3) must discard their signature way of rendering change non-contradictory. But I think that perdurantists should not discard their signature way of rendering change non-contradictory. After all, this way of rendering change non-contradictory is the most

[5] Endurantists deny that persisting objects have temporal parts. So endurantists cannot avail themselves of the perdurantists' way of rendering change non-contradictory. So endurantists need another way. My way is presentism (see Merricks, 1994; 1995; and 2007, 119–25). So I deny that there are any objects that exist only at future times. But then how do I make sense of the claim that a presently existing person x is numerically identical with a person y who exists at a future time? I say this claim is true just in case person x exists and x = y and y will exist at a future time (for details, see Merricks, 1994).

familiar motivation for perdurance (cf. Rea, 1998, 240). Moreover, perdurantists cannot replace this way with just any other way of making change non-contradictory.

For example, presentism is one way of making change non-contradictory (see, e.g. Merricks, 2007, 119–25). But perdurance is not consistent with presentism. This is because perdurance implies that a persisting entity has temporal parts located at multiple times. You can have something as a part only if that something exists. So perdurance requires that your temporal parts located at multiple times all exist. This includes your temporal parts located only at past times and only at future times. But presentism denies that any entities exist that exist only at past times or only at future times, that is, that merely did exist or merely will exist. So perdurance is not consistent with presentism. (See Merricks (1995) for a more detailed version of this argument.)

Another way of making change non-contradictory takes all (allegedly) temporary properties of persisting objects to be time-indexed (or relations to times). Let F be an (alleged) temporary property. Then, according to this attempt, it is not the case that person P is F simpliciter and later that person P will not be F simpliciter. Rather, according to this attempt, it is always the case that P is F-at-t and not-F-at-t*, where t* is later than t. And P can have both of these time-indexed properties without threat of contradiction.

Recall that perdurance implies that the temporal parts of a person have that person's temporary properties (i.e. properties with regard to which that person changes). Let F be such a property. I do not think that anyone would want to say that your instantaneous temporal part at time t is only F-at-t, but not F simpliciter. So I suppose that perdurantists who attempt to make change non-contradictory by time-indexing the properties of persisting objects should say that while a person's temporal part at time t can be F simpliciter, that perduring person is not F simpliciter, but only F-at-t.

Presumably, the perduring person is F-at-t just in case their temporal part at time t is F. But now we are back to a schema considered above: for every person P and every property F, if P has a temporal part at t that is F, then P is F-at-t (in virtue of having a temporal part at t that is F). As already noted, this schema is consistent with the view that only temporal parts have experiences and only temporal parts have

future-directed self-interested concern with regard to experiences. So—as already noted—this schema does not lead to (4D3).

More importantly, (4D3) is not consistent with the view that no person has future-directed self-interested concern or has an experience, but instead has only time-indexed properties that amount to having a temporal part that has future-directed self-interested concern, or a temporal part that has an experience. So time-indexing all of a perduring person's temporary properties rules out (4D3). So the perdurantists we are here considering should deny that time-indexing all temporary properties of a persisting object is an acceptable way to render change non-contradictory. For these perdurantists endorse (4D3).

There is one other problem for perdurantists who attempt to make change non-contradictory by taking all temporary properties of persisting objects to be time-indexed (or relations to times). This problem begins by noting that change with regard to parts is as familiar as is change with regard to properties. So those who seek to avoid change-induced contradiction by taking all temporary properties to be time-indexed should say the same about having parts. That is, they should deny that a perduring object ever has a part simpliciter, but instead only has a part-at-a-time.

The view that parthood is always indexed to a time robs us of the resources to state perdurance itself, and in particular to differentiate it from endurance. Elsewhere (Merricks, 1999b), I defend this claim in detail. Here let me just present the basic idea. That basic idea is that the difference between perdurance and endurance comes down to whether an object's 'parts at some past (or future) time' are thereby that object's parts simpliciter: perdurantists say 'yes', endurantists say 'no'. Perdurantists should want to be able to state perdurance, and in particular to differentiate it from endurance. So perdurantists should not attempt to make change non-contradictory by taking all temporary properties to be time-indexed (or to be relations to times).

Maybe there is some acceptable way for perdurantists to avoid contradiction in the face of change while also endorsing (4D3). If so, then perdurantists might be able to take (3) to be true and to apply to actual cases of future-directed self-interested concern. Then perdurantists might be able to endorse Section I's argument for numerical identity's being a good answer to the Why Question and also take that answer to apply to actual cases of survival. That would be great!

But I do not think that perdurantists should endorse Section I's argument for numerical identity's being a good answer to the Why Question that applies to actual cases of survival. For I have no idea how perdurantists can avoid contradiction in the face of change while also endorsing (4D3). So I conclude that if persons perdure, then what we normally think of as the temporary properties of a person are not really had by that person at all, but are had, instead, only by that person's temporal parts. This conclusion reinstates—and is required to reinstate—the standard perdurantist way of rendering change non-contradictory. So I think this conclusion reinstates standard perdurance.

If persons perdure, then persons do not have temporary properties. So no person has future-directed self-interested concern with regard to an experience. Nor does a person first-personally anticipate an experience. If persons never have future-directed self-interested concern, then you might conclude that it is never appropriate for a person to have such concern. And if persons never first-personally anticipate any experience, then you might conclude it is never appropriate for persons to do so. And so you might conclude that, if persons perdure, then no one will ever have what matters in survival for anyone. That is, you might conclude that perdurance precludes survival.

Perdurantists will want to avoid this conclusion. They can do so by endorsing a perdurantist-friendly adaptation of my answer to:

> *The What Question*: What is it for a person at a future time to have (at that time) what matters in survival for you?

The perdurantist-friendly adaptation: its being appropriate for your temporal part that is located at the present time to first-personally anticipate the experiences that that person's temporal part that is located at that future time will have; and, if that person's temporal part that is located at that future time will have good (or bad) experiences, its being appropriate for your temporal part that is located at the present time to have future-directed self-interested concern with regard to those experiences. Henceforth, let us take perdurantists to answer the What Question in this perdurance-friendly way.

Again, if persons perdure, then persons do not have temporary properties. In other words, if persons perdure, then the only properties that a person has are the properties that that person always has. So if persons perdure, then the only properties that you have are the

properties that you always have. You may always have the determinable property *having some shape or other*. But there is no determinate shape that you always have. More generally, there are no (or hardly any) determinate physical properties that you always have. Nor are there any determinate mental properties that you always have. So if persons perdure, you do not have any (or almost any) determinate physical properties or any determinate mental properties. So if persons perdure, then you are unrecognizable, both to yourself and to others. I myself think this is an absolutely compelling reason to reject perdurance.

This reason might lead someone who starts out as a perdurantist not only to reject perdurance, but also to endorse a third and final view about the metaphysics of the persistence of persons. To see why I say this, suppose you start out believing that you perdure. But then you recognize that, given perdurance, the (determinate) physical and mental properties you take yourself to have are had not by you, but instead are had by your current temporal part. You might then conclude that you are what you originally thought was merely your current temporal part. That is, you might conclude that you are an instantaneous *stage*. Now you are a *stage theorist*.[6]

Stage theory says that each of us is an instantaneous stage. This implies that each person is instantaneous. Nevertheless, stage theorists also maintain that persons persist.[7] In particular, stage theorists say that a person (who is a stage) who is located at time t persists until a later time t* just in case that person (stage) stands in the 'I-relation' to a numerically distinct person (stage) who is located at time t*. Because a stage can be I-related to a numerically distinct stage, the I-relation is

[6] Again, you might go from being a perdurantist to being a stage theorist because you recognize that if you perdure, then you do not have the determinate physical properties or the determinate mental properties that you think you have; whereas you do have those properties if stage theory is true. So I conclude that perdurance and stage theory differ with regard to what properties you have. This conclusion is just one reason that I object when the stage theorists Katherine Hawley (2001, 202 and 208) and Theodore Sider (2001, 191–2) claim that stage theory and perdurance agree on the metaphysics and differ only semantically (for other reasons, see Merricks, 2003).

[7] But I object that if each of us is instantaneous, then we do not persist (see Merricks, 2003).

obviously not numerical identity. As we shall see in Chapter 3 (§III), stage theorists typically take the I-relation to be a psychological relation.[8]

Stage theorists analyze the property *will have experience E* as: stands in the I-relation to a numerically distinct person (stage), located at a later time, who has experience E at that time (cf. Sider, 1996, 437). So the person (stage) who has the property *will have experience E* is not numerically identical with the person (stage) who is located at, and has E at, a later time. So stage theory implies that the expression 'will have experience E' is ambiguous. That expression could mean: has the property *will have experience E*. Or it could mean: is located at, and has experience E at, a later time.

Stage theory implies that 'will have experience E' is ambiguous. So stage theorists will take the following to be ambiguous:

> (3) For all persons x and all persons y, x's being numerically identical with y and the fact that y will have good (or bad) experience E both implies that it is appropriate for x now to have future-directed self-interested concern with regard to E and also explains why this is appropriate.

Stage theorists will disambiguate (3) by disambiguating the expression 'will have good (or bad) experience E' as it occurs in (3).

So stage theorists will take one disambiguation of (3) to be:

> (ST3a) For all persons x and all persons y, x's being numerically identical with y and the fact that y is located at, and has good (or bad) experience E at, a later time both implies that it is appropriate for x now to have future-directed self-interested concern with regard to E and also explains why this is appropriate.

Let x be an instantaneous stage that now has future-directed self-interested concern; then x is located at, and only at, the present time; then x is not located at a later time. Let y be a stage that is located at, and has E at, a later time. Then it is false that x is identical with y. This

[8] The expression 'I-relation' is from Sider (1996). Stage theorists take persistence to be sortal-relative (see Sider, 1996; Hawley, 2001, 156–8). I shall say more about this sortal-relativity in Ch. 3 (§IV). But for now, just note that the stage theorists' account of the persistence of a person described here is their account of the persistence of a person *qua person*.

illustrates that stage theorists should conclude that (ST3a) does not tell us what explains the appropriateness of actual cases of future-directed self-interested concern.

Here is the other stage-theoretic disambiguation of (3):

> (ST3b) For all persons x and all persons y, x's being numerically identical with y and the fact that y has the property *will have good (or bad) experience E* both implies that it is appropriate for x now to have future-directed self-interested concern with regard to E and also explains why this is appropriate.

Stage theorists can say that a person (stage) who has future-directed self-interested concern with regard to a future experience E is numerically identical with a person (stage) who has the property *will have experience E*. For this is just to say that the stage that has future-directed self-interested concern with regard to E is (numerically identical with a stage that is) I-related to a stage located at a future time that has E at that future time. So stage theorists who take (3) to amount to (ST3b) can take (3) to tell us what explains the appropriateness of actual cases of future-directed self-interested concern.

As we saw in Section I, endurantists should take (3) to lead to the conclusion that being numerically identical with is a good answer to:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

But stage theorists who endorse (3) (understood as (ST3b)) should instead take (3) to lead to the conclusion that your being I-related to a person at a future time is a good answer to the Why Question. This is because stage theorists say that a person (stage) has the property *will have E* just in case that person (stage) is I-related to a person (stage) located at a future time who has E at that future time.

Consider the following three claims: the relation of 'identity over time' is numerical identity; persons are the relata of identity over time in cases of personal identity over time; and persons have future-directed self-interested concern with regard to experiences, and have experiences. The argument in Section I for my answer to the Why Question both succeeds and also applies to actual cases of survival only if all three of these claims are true.

Endurantists accept all three claims. But stage theorists reject the first claim. And perdurantists should reject the third claim, as I have argued above. So I think that endurantists—and only endurantists—should both endorse my argument for numerical identity's being a good answer to the Why Question, and also take that answer to apply to actual cases of survival. So, for the rest of this book, I shall assume that my argument for my answer to the Why Question, along with that answer's applying to actual cases of survival, requires endurance.

Perhaps you endorse some fourth metaphysics of persistence, not here (or anywhere) examined, that does not imply endurance but does imply the above three claims. Or perhaps you endorse some non-standard version of perdurance that implies the above three claims without leading to contradiction in the face of change. Then, unlike me, you should conclude that Section I's argument for answering the Why Question in terms of numerical identity does not require endurance. So that argument should be endorsed by more than just endurantists. That would be great!

III. Not the Criterion of Personal Identity over Time

A c*riterion of personal identity over time* is supposed to be necessary and sufficient for personal identity over time, but not presuppose personal identity over time. According to John Locke, having 'the same consciousness' is the criterion of personal identity over time. And Locke seems to think that its being appropriate for you to have future-directed self-interested concern for a person's experiences at a future time is explained by your having the 'same consciousness' as that person at that time. (See Locke, *An Essay Concerning Human Understanding*, II, xxvii, §§15–19 [1975, 340–2].)

So Locke's view seems to be that your surviving as a person at a future time is explained by your having 'the same consciousness' as that person at that time. That is, Locke's view seems to be that your being related to a person at a future time by (what Locke takes to be) the criterion of personal identity over time explains why you will survive as that person at that time. Thus Locke seems to accept:

The Criterion Explains Survival: If your being the same person as a person who will exist (and be conscious) at a future time is sufficient

> for that person to have, at that time, what matters in survival for you, then your being related to that (conscious) person by the criterion of personal identity over time explains why that person will have, at that time, what matters in survival for you.

The Criterion Explains Survival is widely accepted nowadays. In particular, the Criterion Explains Survival seems to be accepted by almost every philosopher who takes personal identity over time to be sufficient for survival. And almost all these philosophers endorse a more or less Lockean criterion of personal identity over time, that is, a psychological criterion of personal identity over time.

For example, Derek Parfit (1984) endorses the following three claims: personal identity over time is sufficient for survival; unbranching psychological connectedness and/or continuity is the criterion of personal identity over time; and psychological connectedness and/or continuity—including unbranching psychological connectedness and/or continuity—explains survival (see Ch. 3, §II).

Moreover, the Criterion Explains Survival has played a key role in motivating a psychological criterion of personal identity over time. For example, David Shoemaker (2016) and also Jens Johansson (2016) independently argue as follows. Being identical with a person who exists at a future time is sufficient for that person's having, at that time, what matters in survival for you. This implies that your being related to that person by the criterion of personal identity over time explains why that person will have what matters in survival for you. (To endorse this implication is to endorse the Criterion Explains Survival.) Your being biologically continuous with a person at a future time would not explain why that person will have what matters in survival for you. So biological continuity is not the criterion of personal identity over time. Rather, the correct criterion of personal identity over time is psychological continuity.

Bernard Williams's 'The Self and The Future' (1970) takes personal identity to be sufficient for what matters in survival. And Williams (1970) favors the view that your being 'bodily continuous' with a person at a future time explains why that person will have what matters in survival for you. And he takes this view to imply that you enjoy personal identity with a person at a future time just in case that person is bodily continuous with you. So Williams (1970) might be the

extremely rare example of someone taking the Criterion Explains Survival to motivate a criterion of personal identity over time that is not psychological.[9]

The Criterion Explains Survival has also been presupposed in an argument for the conclusion that personal identity over time is not sufficient for survival. For Eric Olson (1997, 52–72) argues as follows. Biological continuity is the criterion of personal identity over time. Biological continuity does not explain survival. This implies that it is false that your being identical with a future person is sufficient for your surviving as that person. (To endorse this implication is to endorse the Criterion Explains Survival.)

I have argued that numerical identity is a good answer to:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

The Criterion Explains Survival is widely assumed. Those who assume the Criterion Explains Survival might misinterpret my answer to the Why Question. That is, they might think that I am really answering the Why Question in terms of the criterion of personal identity over time, as opposed to in terms of numerical identity itself.

So let me be clear. I answer the Why Question in terms of numerical identity itself. Moreover, the argument in Section I is an argument for the conclusion that being numerically identical with is a good answer to Why Question. The argument in Section I is not an argument for the conclusion that being related by the criterion of personal identity over time is a good answer to the Why Question. I have not endorsed—and have given no argument for—the claim that

[9] Or perhaps not. For Williams (1970) seems to understand 'bodily continuity' as having the numerically same body. And this might not count as a genuine criterion of personal identity over time. For suppose that having numerically the same body is necessary and sufficient for personal identity over time only if a person is a body. Then a future person's having the numerically same body as you just is that future person's being the numerically same body as you. This presupposes that that person is (a body and is) numerically identical with you. But a genuine criterion of personal identity over time would not presuppose that that person is numerically identical with you.

there is a good answer to the Why Question in terms of the criterion of personal identity over time.

Numerical identity's being a good answer to the Why Question is consistent with there being other good answers, including an answer in terms of the criterion of personal identity over time. But since we already have a good answer in terms of numerical identity, I myself see no reason to add that there is another good answer to the Why Question in terms of the criterion of personal identity over time. So I do not endorse the Criterion Explains Survival.[10] So I am not persuaded by the above arguments against the biological criterion of personal identity from Shoemaker and from Johansson, or by Olson's argument above against the sufficiency of personal identity over time for what matters in survival.

I answer the Why Question in terms of numerical identity. Again, I see no reason to add that there is another good answer to the Why Question in terms of the criterion of personal identity over time. So I see no reason to endorse the Criterion Explains Survival. But you might object that there is a reason for me in particular to endorse the Criterion Explains Survival. For you might object that numerical identity's being a good answer to the Why Question leads to the conclusion that another good answer to the Why Question is in terms of the criterion of personal identity over time. If you think this, it is probably because you endorse one of two arguments.

Here is the first argument. Numerical identity is a good answer to the Why Question. You persist by enduring. So your surviving as a future (conscious) person is explained by your being the same person as that (conscious) person. Being the same person as a person at another time is 'reduced to' being related to that person at another time by the criterion of personal identity over time. So personal identity over time is 'reduced to' being related to a person at another time by the criterion

[10] I think that the Criterion Explains Survival is false. I think that it is false because I endorse its antecedent and reject its consequent. For I take identity (with a conscious person) to be sufficient for survival; but I deny that there are any criteria of identity over time at all; a fortiori, I deny that the criterion of personal identity explains survival. I shall not rely on this denial in the text. But for my reasons for denying that there are any criteria of identity over time, see Merricks, 1998 and 2001b, 191–5.

of personal identity over time. Whatever personal identity over time explains is thereby explained by what personal identity over time is reduced to.[11] Therefore, your being related by the criterion of personal identity over time to a (conscious) person at a future time explains why you will survive as that person at that time.

I object to this argument because I think that it is false that personal identity over time is reduced to being related to a person at another time by the criterion of personal identity over time. One reason that I object in this way begins with the observation that a person can never differ from herself. So a person has exactly the same psychological states as that person has. Psychological continuity is not supposed to be the same thing as having exactly the same psychological states. So I conclude that the relata of psychological continuity are not persons themselves, but are instead the way a person is psychologically at one time and the way a person is at another time. For a person can be psychologically one way at one time, and psychologically a different way at a different time.

I do sometimes say things like: 'You are psychologically continuous with a person at a future time.' But I take this to be shorthand for the claim that the way you are now is psychologically continuous with the way a person will be at a future time. Other endurantists should join me. Perdurantists should also take 'you are psychologically continuous with a person at a future time' to be shorthand. But, unlike endurantists, perdurantists should take this to be shorthand for the claim that your current temporal part is psychologically continuous with a future temporal part of a person (see, e.g. Lewis, 1976). Only stage theorists can take you (they think you are a stage) to be quite literally psychologically continuous with a person (a stage) at a future time.

I say that persons endure. So I say that the relata of psychological continuity are the ways a person is at a time. The same goes for the relata of psychological connectedness. And the same goes for the relata of other sorts of continuity and connectedness. Assume that a criterion of personal identity over time is some sort of continuity or connectedness. Then the relata of the criterion of personal identity over time are the way a person is at one time and the way a person is at another

[11] This premise is controversial. At least, Mark Johnston (1997) controverts it.

time. But the relata of personal identity over time are persons. A person is not the same thing as (and cannot be reduced to) the way a person is at a time. So the relata of the criterion of personal identity over time are not (and cannot be reduced to) the relata of personal identity over time. So I conclude that it is false that personal identity over time is reduced to being related to a person at another time by the criterion of personal identity over time.

Here is another reason to draw the same conclusion, but a reason that has to do with the relation, as opposed to the relata. I think that you persist by enduring. So your now being the same person as someone existing at a future time just is your being a person who is now wholly present and who stands in the relation *being numerically identical with* to a person who will be wholly present at that future time. The relation *being numerically identical with* is not reduced to being psychologically continuous with.[12] So it is false that your being the same person as a person existing at a future time is reduced to your being a person who is now wholly present and who is psychologically continuous with a person who will be wholly present at that future time. So if psychological continuity is the criterion of personal identity over time, it is false that personal identity over time is reduced to being related to a person at another time by the criterion of personal identity over time.[13]

The argument just given focused on psychological continuity. But it should be clear that—if persons endure—a parallel argument can be run for the conclusion that personal identity over time is not reduced to being related to a person at another time by psychological connectedness, or by biological continuity, or by any of the other contenders for being the criterion of personal identity over time. Thus we have my second reason for concluding that personal identity over time cannot be reduced to being related to a person at another time by the criterion of personal identity over time.

[12] Everything is numerically identical with itself. But not everything is psychologically continuous with itself. This is just one reason to conclude that the relation *being numerically identical with* is not reduced to being psychologically continuous with.

[13] For discussion of what it is to be 'wholly present', see Merricks (1994; 1999a; 1999b). For a more detailed version of the argument just given, including replies to some objections, see Merricks (1999a).

For the above two reasons, I conclude that if persons endure, it is false that personal identity over time is reduced to—or just is—being related to a person at another time by the criterion of personal identity over time. In other words, if persons endure, then personal identity over time involves a 'further fact' distinct from being related to a person at another time by the criterion of personal identity over time. In particular, personal identity over time involves the 'further fact' of being related by numerical identity.

Parfit (1984) seems to think that our being 'Cartesian egos' is the only way to avoid the result that personal identity over time just is being related by unbranching psychological connectedness and/or continuity. But we have just seen that our persisting by enduring is another way to avoid this result. We shall return to this point in Chapter 3 (§IV). For this point plays a role in Chapter 3's critique of Parfit's argument for the conclusion that personal identity over time is not necessary for survival.

We have just considered one argument for the claim that numerical identity's explaining survival leads to the conclusion that the criterion of personal identity also explains survival. That argument turned on the thesis that being the same person as a person at a future time is reduced to being related to that person by the criterion of personal identity over time. But we have seen that if persons endure, then that thesis is false. I think that persons endure. So I think we should reject the above argument for the claim that numerical identity's explaining survival leads to the conclusion that the criterion of personal identity also explains survival. Let us now turn to another argument for this same claim.

Recall that the criterion of personal identity over time is supposed to be (necessary and) sufficient for personal identity over time without presupposing personal identity over time. This suggests that the criterion of personal identity over time explains personal identity over time. So—if persons endure—this suggests that the way a person is at an earlier time being related by the criterion of personal identity to the way a person is at a later time explains why the person at the earlier time is numerically identical with the person at the later time. Using the sort of shorthand noted above, we can put this suggestion as:

(I) Your being related by the criterion of personal identity over time to a (conscious) person at a future time explains why you are numerically identical with that (conscious) person at that time.

Add:

> (II) Your being numerically identical with a (conscious) person at a future time explains why you will survive as that person at that time.

Then conclude:

> (III) Your being related by the criterion of personal identity over time to a (conscious) person at a future time explains why you will survive as that person at that time.

I do not think this argument is sound.[14] But suppose that it is sound. Then premise (II) is true. And if premise (II) is true, then numerical identity is a good answer to the Why Question. Moreover, if persons endure, then your being related to a person at a future time by the criterion of personal identity over time implies that you are numerically identical with that person. So this argument's conclusion is consistent with my claim—defended in Chapter 3 (§I)—that every good answer to the Why Question implies numerical identity. So even though I do not think the above argument is sound, I do not need to oppose it. For the above argument is no threat to anything I say about the Why Question.

Suppose that you take the above argument to be sound. Then you will conclude that the Criterion Explains Survival is true. Then you cannot fault the reasoning at the start of this section from Shoemaker, Johansson, and Olson for presupposing the Criterion Explains Survival. But you should nevertheless think that their reasoning is faulty.

For if you take the above argument to be sound, then you should conclude that the criterion of personal identity over time explains survival no matter what that criterion turns out to be. That criterion explains survival—according to the above argument—by explaining personal identity over time, which in turn explains survival. So you should deny Olson's (1997) claim that biological continuity is the criterion of personal identity but does not explain survival. Moreover, you should conclude that Shoemaker's (2016) and Johansson's (2016) claim that biological continuity does not explain

[14] I think this argument's first premise—(I)—is false; this is because I deny that there is a criterion of personal identity over time. Some might say that this argument is not valid, perhaps because they deny that explanation is transitive.

survival presupposes (and so cannot be used to argue for) the claim that biological continuity is not the criterion of personal identity over time.

Again, suppose that you endorse the Criterion Explains Survival because of the above argument. Then you should not only object to Olson (1997), Shoemaker (2016), and Johansson (2016). You should also object to the idea that an answer to Why Question in terms of the criterion of personal identity over time would be 'deeper' than such an answer in terms of numerical identity. You should conclude, instead, that the 'deeper' answer is in terms of numerical identity itself. For—according to the above argument—being related by the criterion of personal identity does not explain survival directly, but instead explains survival only by way of explaining the holding of numerical identity. It is the holding of numerical identity itself that directly explains survival.

IV. An Unanswered Question

Recall:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

To answer this question is to explain why it is appropriate for you to first-personally anticipate, and have future-directed self-interested concern with regard to, the experiences that a person will have at a future time. So any answer to this question presupposes that there will be a person who will have the experiences in question. So any answer to this question presupposes that those experiences will be had by someone.

Let me illustrate this with Jones and the Nobel Prize in philosophy. Leibniz would claim that Jones's first-personally anticipating, and having future-directed self-interested concern with regard to, receiving the Prize is appropriate if the person who will receive the Prize next week is not only Jones, but also will then have memories of Jones's life today (Ch. 1, §I). Whiting would claim that this is appropriate if the person who will receive the Prize next week will then have the values, desires, and projects that Jones now has (see Ch. 4, §I). Velleman would claim that this is appropriate if Jones has 'first-personal access' to the point of view that a person will have when receiving the Prize next week (see Ch. 4, §II). I claim that this is

appropriate if Jones is numerically identical with the person who will receive the Prize next week. And so on. All of these claims presuppose that a person will receive the Prize next week.

Here is a new version of the story of Jones and the Prize. Jones has won the Prize. And Jones is slated to receive the Prize at a ceremony next week. Jones knows all this. And she is right now first-personally anticipating receiving the Prize. But she will die in a tragic accident tomorrow. So she will not receive the Prize next week. Indeed, no one will receive the Prize next week. Is it now appropriate for Jones to first-personally anticipate receiving the Prize next week? No answer to the Why Question answers this question. For this is a question about an experience that no one will have. Moreover, nothing defended in this book answers this question. For this book is concerned with what matters in survival. And if no one will receive the Prize next week, then it is false that the person who will receive the Prize next week will then have what matters in survival for Jones. Whether or not it is now appropriate for doomed Jones to first-personally anticipate receiving the Prize is neither here nor there.

I suppose that an answer to the Why Question might inspire an answer to: is it now appropriate for doomed Jones to first-personally anticipate receiving the Prize next week? For example, some who endorse my answer to the Why Question might claim that Jones's first-personally anticipating receiving the Prize next week is appropriate because, had Jones not died, someone numerically identical with Jones would have received the Prize next week. And some who endorse Leibniz's answer to the Why Question might claim that this is appropriate because, had Jones not died, Jones would have received the Prize next week while enjoying memories of her earlier life. And so on. But this is all just speculation and has no bearing on what follows.

V. Conclusion

I have argued in this chapter that being numerically identical with is a good answer to:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

My argument for that answer turned on a particular metaphysics of persistence, namely, endurance (§I). Suppose that persons endure. Do all good answers to the Why Question imply being numerically identical with? Or suppose that persons do not endure. Do all good answers to the Why Question imply personal identity over time, that is, being the same person as, or persisting as, that person at that future time? These are questions about the necessity of personal identity for what matters in survival. These questions take center stage in Chapter 3.

3

On the Necessity of Personal Identity

Joseph Butler (1736 [1975]) claims that Locke's views on personal identity imply that your self-interested concern with regard to a person's experiences at a future time can be appropriate even if you will not be identical with that person. Butler then objects to Locke's views on account of this implication. Henry Sidgwick (1907, 419) raises a similar objection to Hume's views on personal identity. These objections reveal that Butler and Sidgwick take personal identity over time to be necessary for what matters in survival. So too do some of the philosophers discussed in the preceding chapters.[1] And so do I. In this chapter, I shall say why I take personal identity to be necessary for what matters in survival (§I). I shall also show how Derek Parfit's famous argument to the contrary fails (§§II–IV).

I. An Argument for the Necessity of Personal Identity

Chapter 2 began by observing that it is appropriate for you now to have self-interested concern with regard to pain that you are now experiencing. This chapter begins by observing that it is not appropriate for you

[1] I read Leibniz as taking personal identity to be necessary (but not sufficient) for what matters in survival (see Ch. 1, §§I–II). McMahan (2002, 79) also takes personal identity to be thus necessary. And consider that Schechtman says: 'reidentifying persons . . . constrains (but does not determine) the kind of psychological configurations that can constitute a single psychological subject' (1996, 69) and 'For a person to have *self-interested* concern for the fulfillment of a desire, that desire must, to be sure, at least be part of his history' (1996, 85).

Self and Identity. Trenton Merricks, Oxford University Press. © Trenton Merricks 2022.
DOI: 10.1093/oso/9780192843432.003.0004

now to have self-interested concern with regard to pain that some other person (but not you) is now experiencing. More generally:

> (1*) For all persons x and all persons y, if x is numerically distinct from y and y (but not x) has experience E at the present time, then it is not appropriate for x now to have self-interested concern with regard to E.

I think that (1*) is true. Moreover, I think parallel claims are true about what was and what will be. For example:

> (2*) For all persons x and all persons y, if x is numerically distinct from y and y (but not x) will have experience E at a future time, then it will not be appropriate at that future time for x to have self-interested concern with regard to E.

I think that (1*) and (2*) illustrate that a person x's being numerically distinct from a person y is sufficient for it not to be appropriate for x to have self-interested concern with regard to y's present or future experiences. So I further conclude:

> (3*) For all persons x and all persons y, if x is numerically distinct from y and y (but not x) will have experience E, then it is not appropriate for x now to have future-directed self-interested concern with regard to E.

A person existing at a future time will have, at that time, what matters in survival for you only if it is now appropriate for you to have future-directed self-interested concern with regard to that person's good (or bad) experiences at that time (Ch. 1, §II). So it should be obvious that (3*) is somehow relevant to who will (and who will not) have what matters in survival for you. But the precise way in which (3*) is thus relevant is less obvious. For the precise way depends on the metaphysics of persistence.

For starters, suppose that you persist by enduring. Then you are numerically identical with a person who will have experience E just in case, first, it will be the case that a person has E and, second, you are numerically identical with that person. So (3*) implies that if a person at a future time is numerically distinct from you, then it is not appropriate for you to have future-directed self-interested concern with regard to that person's good (or bad) experiences at that time.

So if a person at a future time is numerically distinct from you, then that person will not have what matters in survival for you.

In other words, a person at a future time will have what matters in survival for you only if you are numerically identical with that person. So every good answer to the following implies being numerically identical with:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

So numerical identity is necessary for survival. So—given how endurantists understand personal identity—personal identity is necessary for survival. That is, a person at a future time will have what matters in survival for you only if you are the same person as that person. None of this seems to be contingent. So—being an endurantist—I further conclude that, necessarily, a person at a future time will have what matters in survival for you only if you are the same person as that person.[2]

Now turn to stage theory. Recall from Chapter 2 (§II) that the expression 'will have experience E' is ambiguous, given stage theory.

[2] I deny that two persons can have the numerically same token conscious experience E. So I take '(but not x)' in (1*) through (3*)—which I include for emphasis—to be strictly speaking redundant. But suppose that the following case is possible: a person at a future time is numerically identical with you, and that person's (that is, your) conscious experiences at that time are numerically identical with a numerically distinct person's conscious experiences at that time. Then the '(but not x)' in (3*) is not redundant. And then the combination of (3*) and endurance is consistent with (but does not imply) the claim that—in the case just considered—it is appropriate for you to have future-directed self-interested concern with regard to the experiences that are had by each of those persons at that future time. This claim implies that someone who is not numerically identical with you will have what matters in survival for you. But even given this claim, there is still a way in which numerical identity is necessary for survival. For even given this claim, you will survive at a future time as a person who is not numerically identical with you only if—and only because—you also survive at that time as a person who is numerically identical with you (and both persons have, at that time, the numerically same conscious experiences). This is definitely not the sort of survival without identity that Parfit defends. That is one reason that I shall ignore the above case in what follows. But the main reason is that I think this case is not possible; for, again, I deny that two persons can have the numerically same token conscious experience E.

That expression could mean: is located at, and has experience E at, a later time. Or it could mean: has the property *will have experience E*. As a result, (3*) itself is ambiguous, given stage theory. So stage theorists should not ask what (3*) implies with regard to who will have what matters in survival for you. Rather, stage theorists should disambiguate (3*), and then ask what each disambiguation implies with regard to who will have what matters in survival for you.

Suppose that stage theorists take 'will have experience E' in (3*) to mean: is located at, and has experience E at, a later time. Then they should take (3*) to amount to:

> (ST3*a) For all persons x and all persons y, if x is numerically distinct from y and y (but not x) is located at, and has experience E at, a later time, then it is not appropriate for x now to have future-directed self-interested concern with regard to E.

Let x be an instantaneous stage (person) that now has future-directed self-interested concern with regard to E; then x is located at, and only at, the present time. Add that y is located at, and has experience E at, a later time; then y is located at a later time. So it is false that x is identical with y. This illustrates that if (ST3*a) is true, it is never appropriate for a person (stage) to have future-directed self-interested concern with regard to a future experience. So if (ST3*a) is true, stage theory precludes survival. I assume that stage theorists will reject (ST3*a).

Suppose that stage theorists take 'will have experience E' in (3*) to mean: has the property *will have experience E*. Then they should take (3*) to amount to:

> (ST3*b) For all persons x and all persons y, if x is numerically distinct from y and y (but not x) has the property *will have experience E*, then it is not appropriate for x now to have future-directed self-interested concern with regard to E.

Let x be a person (stage) who now has future-directed self-interested concern with regard to E; then x is located at the present time. Let z be a person (stage) who is located at a later time and has E. Stage theory allows for x to be I-related to z. To have the property *will have experience E* just is to be I-related to a stage that is located at, and has experience E at, a later time (Ch. 2, §II). So stage theory allows for

x to have the property *will have experience E*. So stage theory allows for x to be numerically identical with the person (stage) who has the property *will have experience E*. This illustrates that (ST3*b) does not imply that stage theory precludes survival.

Above, I motivated (3*). I think that that motivation is compelling. And I think it should be compelling to stage theorists. For example, stage theorists should accept that motivation's starting point:

> (1*) For all persons x and all persons y, if x is numerically distinct from y and y (but not x) has experience E at the present time, then it is not appropriate for x now to have self-interested concern with regard to E.

So I think that stage theorists should accept (3*). More carefully, I think stage theorists should accept at least one of (what they take to be) the disambiguations of (3*). They will reject (ST3*a). So I think they should accept (ST3*b). So let us assume that stage theorists accept (3*), understood as (ST3*b). Then stage theorists should conclude that it is appropriate for you (a stage) to have future-directed self-interested concern with regard to the experiences of a person (stage) who is located at a later time only if you are I-related to that person (stage).

So stage theorists should conclude that you survive as a person (stage) located at a future time only if you are I-related to that person (stage). This conclusion does not have the result that every good answer to the Why Question implies being numerically identical with. Rather, it has the result that every good answer to the Why Question implies being I-related to. So—given how stage theorists will understand personal identity—stage theorists should conclude that personal identity is necessary for survival. That is, they should conclude that, necessarily, a person at a future time has what matters in survival for you only if you are the same person as that person.

Now turn to perdurance. Suppose that perduring persons have 'temporary' properties, including having future-directed self-interested concern with regard to an experience, and having that experience. Then perdurantists who endorse (3*) should claim that it is appropriate for a perduring person to have future-directed self-interested concern with regard to an experience E only if that numerically same perduring person has experience E. This claim implies that perduring

is necessary for survival. So perdurantists who endorse this claim should conclude that persisting is necessary for survival. So—given how perdurantists understand personal identity—they should conclude that personal identity is necessary for survival.

But I think that perdurantists should deny that persons have 'temporary' properties. I think that perdurantists should say that those properties are had instead only by temporal parts (Ch. 2, §II). Suppose I am right. Then perdurantists should deny that persons have future-directed self-interested concern with regard to an experience, or have that experience. So perdurantists should conclude that any claim about 'all persons' who have future-directed self-interested concern, or who have an experience, is a claim that has no instances. So perdurantists should say that the following is a claim that has no instances:

> (3*) For all persons x and all persons y, if x is numerically distinct from y and y (but not x) will have experience E, then it is not appropriate for x now to have future-directed self-interested concern with regard to E.

So perdurantists should say that even if (3*) is true, (3*) is irrelevant to actual cases of survival. So perdurantists can ignore (3*).

But (3*) might suggest the following:

> (3**) For all x and all y, if x is numerically distinct from y and y (but not x) will have experience E, then it is not appropriate for x now to have future-directed self-interested concern with regard to E.

Unlike (3*), (3**) quantifies over everything, not just over persons. So if there are temporal parts, (3**) quantifies over temporal parts, among other things. If persons perdure, then future-directed self-interested concern with regard to an experience E is had by temporal parts alone, all of which are numerically distinct from any later temporal part that has experience E. So the conjunction of (3**) and perdurance implies that it is never appropriate for any temporal part to have future-directed self-interested concern with regard to an experience. So if (3**) is true, then perdurance has the result that no future person has what matters in survival for anyone.

That perdurance has this result is reinforced by the perdurantist-friendly adaptation (Ch. 2, §II) of my answer to:

> *The What Question*: What is it for a person at a future time to have (at that time) what matters in survival for you?

That adaptation: its being appropriate for your temporal part that is located at the present time to first-personally anticipate the experiences that that person's temporal part that is located at that future time will have; and if that person's temporal part that is located at that future time will have good (or bad) experiences, its being appropriate for your temporal part that is located at the present time to have future-directed self-interested concern with regard to those experiences.[3]

Perdurantists will want to deny that perdurance precludes survival. So they will want to deny (3**). We got to (3**) by way of:

> (3*) For all persons x and all persons y, if x is numerically distinct from y and y (but not x) will have experience E, then it is not appropriate for x now to have future-directed self-interested concern with regard to E.

That is, I said that (3*) might suggest (3**). But perdurantists could reply that (3*) does not suggest (3**). They could say instead that (3*) suggests that in cases of appropriate future-directed self-interested concern with regard to a later experience, only one person should be 'involved' with both the concern and also the experience. One way for one person to be 'involved' with both is for that one person to have both. But a second way—perdurantists might add—is for one person to have as parts both the earlier temporal part that has the future-directed self-interested concern and also the later temporal part that has the relevant experience.

[3] Perdurantists should not respond to (3**) by analyzing *will have experience E* as: is a temporal part of a person who has a later temporal part that has experience E. This is because an entity that exists right now will have an experience only if that entity persists into the future. In particular, a temporal part of a person will have an experience that each later temporal part of that person has only if that temporal part persists for as long as that person persists. But perdurantists should deny that each temporal part of a person persists for as long as that person persists. The parallel to stage theory—which inspires the analysis of *will have experience E* just suggested and discarded—is instructive. Stage theorists think that a person (stage) can have the property *will have experience E* (analyzed in terms of the I-relation) precisely because that stage is a persisting person (and persistence is analyzed in terms of the I-relation).

So perdurantists might take (3*) to lead to:

(3+*) For all x and all y, if x and y are not parts of the same person and y (but not x) has experience E, then it is not appropriate for x now to have future-directed self-interested concern with regard to E.

Perdurantists who endorse (3+*) should conclude that it is appropriate for a temporal part to have future-directed self-interested concern with regard to an experience E only if that temporal part and the later temporal part that has E are parts of the same person. So these perdurantists should conclude that you survive as a person at a future time only if you are the same perduring person as that person. So these perdurantists should take personal identity to be necessary for survival. That is, necessarily, a person at a future time will have what matters in survival for you only if you are the same person as that person.

This section started by defending:

(3*) For all persons x and all persons y, if x is numerically distinct from y and y (but not x) will have experience E, then it is not appropriate for x now to have future-directed self-interested concern with regard to E.

We have seen that (3*) cannot be directly applied to actual cases of future-directed self-interested concern if persons persist stage theoretically, or if persons persist by perduring. Nevertheless, stage theorists should endorse (what they take to be) a disambiguation of (3*); that disambiguation implies that being the same person as a person at a later time as understood by stage theorists is necessary for survival. Similarly, (3*) might motivate perdurantists to endorse a thesis that implies that, necessarily, you survive as a person at a future time only if you are the same perduring person as that person.

If persons persist by enduring, (3*) can be directly applied to actual cases of future-directed self-interested concern. And if persons endure, (3*) implies that you will survive as a person who will exist at a future time only if you are numerically identical with that person. I endorse (3*) and I think that persons endure. So I conclude that being the same person as a person at a later time—that is, personal identity—as understood by endurantists is necessary for what matters in survival.

II. Parfit's Argument against the Necessity of Personal Identity

As we saw in Chapter 1 (§IV), Derek Parfit says:

> An emotion or attitude can be criticized for resting on a false belief or for being inconsistent. A man who regards [double brain hemisphere transplant] as death must, I suggest, be open to one of these criticisms. (1971, 9)

Parfit also says:

> Some people would regard division as being as bad, or nearly as bad, as ordinary death. This reaction is irrational. We ought to regard division as being about as good as ordinary survival. (1984, 261)

These remarks—and the contexts in which they occur—suggest that Parfit takes what matters in survival to be opposed to the practical equivalent of death. The practical equivalent of death—like the practical equivalent of annihilation (Ch. 1, §I)—precludes appropriate future-directed self-interested concern. So we might read Parfit as taking what matters in survival to involve appropriate future-directed self-interested concern. In particular, we might read Parfit as answering the following question in terms of appropriate future-directed self-interested concern:

> *The What Question*: What is it for a person at a future time to have (at that time) what matters in survival for you?

This section focuses on Parfit's influential argument for personal identity's not being necessary for what matters in survival. At least for our purposes, Parfit's argument is most interesting if it is an argument against the conclusion of Section I, namely, that personal identity is necessary for (what I call) what matters in survival. Parfit's argument is an argument against that conclusion if we take him to answer the What Question in terms of appropriate future-directed self-interested concern. And as we saw in Chapter 1 (§II), this is just how Jeff McMahan and Sydney Shoemaker—and, of course, many others—do take Parfit.

But Parfit's argument might not be an argument against that conclusion if we take him to answer the What Question in a different way. And Parfit does sometimes seem to answer the What Question in a different way. Recall from Chapter 1 (§IV) that Parfit makes some

claims using the expression 'what matters in survival' that are claims about what matters to you with regard to the future, such as (his example) being freed of your nicotine addiction at some future time (Parfit, 1984, 299). Suppose we took Parfit to answer the What Question in terms of what matters to you with regard to the future. Then Parfit's conclusion that personal identity is not necessary for (what he calls) what matters in survival might not imply that personal identity is not necessary for (what I call) what matters in survival. This is because personal identity's not being necessary for what matters to you with regard to the future might not imply that personal identity is not necessary for appropriate future-directed self-interested concern or appropriate first-personal anticipation.

Parfit can also seem to answer the What Question in terms of psychological connectedness and/or continuity. Parfit says:

> I said earlier that what matters in survival could be provisionally referred to as 'psychological continuity'. I must now distinguish this relation from another, which I shall call 'psychological connectedness' . . . 'Psychological connectedness' . . . requires the holding of these direct psychological relations . . . 'Psychological continuity' . . . only requires overlapping chains of direct psychological relations . . . Now that we have distinguished the general relations of psychological continuity and psychological connectedness, I suggest that connectedness is a more important element in survival.
>
> (1971, 20–1; see also Parfit, 1984, 262)

Suppose Parfit answers the What Question in terms of psychological connectedness and/or continuity. Then it is trivial that psychological connectedness and/or continuity is sufficient for (what Parfit takes to be) what matters in survival. This triviality is not relevant to whether personal identity is necessary for appropriate future-directed self-interested concern or appropriate first-personal anticipation. So this triviality is not relevant to the conclusion—as defended and understood in Section I—that personal identity is necessary for what matters in survival.

I do not claim that Parfit unequivocally answers the What Question in terms of appropriate future-directed self-interested concern. But I shall assume for the rest of this section that the most common reading of Parfit is correct, and that he does answer the What Question in only this way. This is because, again, this section focuses on Parfit's argument for the claim that personal identity is not

necessary for what matters in survival. That argument is most interesting for our purposes if it clearly opposes the conclusion of Section I, namely, that personal identity is necessary for what matters in survival. Parfit's argument clearly opposes that conclusion if Parfit answers the What Question in terms of future-directed self-interested concern. But Parfit's argument does not clearly oppose that conclusion if Parfit answers the What Question in either of the other two ways just considered.

Let us consider how Parfit answers:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

McMahan says:

> Let us call the relations that ground rational egoistic concern about the future the *prudential unity relations.* According to the traditional view, there is only one such relation: identity. According to Parfit, the prudential unity relations are psychological connectedness and psychological continuity. (2002, 42)

And here is how Shoemaker's critical notice of *Reasons and Persons* describes Parfit's views:

> What makes it rational (to the extent that it is) for me to have a special concern for my well-being at a future time is the fact that my present states stand to my future states in the relations of psychological C&C [connectedness and continuity] . . . the holding of these relations justifies future concern . . . . (1985, 444)

So McMahan and Shoemaker take Parfit to answer the Why Question in terms of psychological connectedness and/or continuity. So do many others. And I agree that Parfit sometimes answers the Why Question in just this way. But some of what Parfit says undermines this answer to the Why Question.

Parfit takes himself to be a 'Reductionist' and says:

> I do not *anticipate* the pain that will be felt by someone I love. It might be claimed that only on the Non-Reductionist View can we justifiably anticipate future pains. Anticipation might be justified only by the non-existent deep further fact. Perhaps if we are Reductionists, we should cease to anticipate our own future pains. (1984, 312)

So Parfit is open to the idea that first-personal anticipation is never appropriate.[4]

If first-personally anticipating an experience is not appropriate, then it is not appropriate to have future-directed self-interested concern with regard to that experience (Ch. 1, §II). And Parfit (1984, 307–10) is also open to the idea that future-directed self-interested concern is never appropriate. This is a familiar point about Parfit. Hence Shoemaker's parenthetical 'to the extent that it is'.

Parfit is open to the idea that neither first-personal anticipation nor future-directed self-interested concern is ever appropriate. This idea implies that there is no good answer to the Why Question. And if there is no good answer to the Why Question, then psychological connectedness and/or continuity is not a good answer to the Why Question. This is why I said above that some of what Parfit says undermines this answer to the Why Question. And to the extent that Parfit undermines this answer to the Why Question, he undermines his argument for the conclusion that personal identity is not necessary for what matters in survival. For, as we shall see, that argument has this premise:

> (III) If you are psychologically connected to and/or continuous with a person at a future time, then that person has, at that time, what matters in survival for you.

There is another way in which the idea that neither first-personal anticipation nor future-directed self-interested concern is ever appropriate undermines Parfit's argument. This idea implies that no one existing at a future time can have what matters in survival for anyone who exists now. So it is false that there can be cases of survival. So it is false that there can be cases of survival without personal identity. So any argument is unsound that concludes that there can be cases of

[4] What is Parfit's Reductionism, and how does it undermine appropriate first-personal anticipation? It is hard to say. Elsewhere (Merricks, 1999c), I look at how a variety of philosophers understand both Parfit's Reductionism and also how Parfit takes Reductionism to bear on what matters in survival. Those philosophers disagree about what Reductionism is, and also about how Parfit takes Reductionism to bear on what matters in survival.

survival without personal identity. So Parfit's argument for that conclusion is unsound.

I do not claim that Parfit unequivocally answers the Why Question in terms of psychological connectedness and/or continuity. Instead, I shall just assume for the rest of this section that he does answer the Why Question in this way. This is a concession to Parfit's argument for the conclusion that there can be cases of survival without personal identity. For Parfit's argument is sound only if there is a good answer to the Why Question and that answer is in terms of psychological connectedness and/or continuity.

Parfit's (1984, 263) argument starts with the claim that personal identity is sufficient for survival. That is, his argument starts with the claim that if a person at a future time is the same person as you, then that person will have, at that future time, what matters in survival for you. Parfit (1984, 262–3) also claims that 'unbranching' psychological connectedness and/or continuity is the criterion of personal identity over time. These two claims imply:

> (I) If you are psychologically connected to and/or continuous with a person—but only that person—at a future time, then that person has, at that time, what matters in survival for you.

Parfit thinks that if unbranching psychological connectedness and/or continuity delivers what matters in survival, then psychological connectedness and/or continuity delivers what matters in survival, regardless of whether that connectedness and/or continuity branches. So Parfit endorses:

> (II) If premise (I) is true, then: if you are psychologically connected to and/or continuous with a person at a future time, then that person has, at that time, what matters in survival for you.

And (I) and (II) imply:

> (III) If you are psychologically connected to and/or continuous with a person at a future time, then that person has, at that time, what matters in survival for you.

Parfit should not say that (III) has the result that you survive as a person at a future time if you are psychologically connected to that person solely in virtue of each of you having the desire to be happy.

So let us take the sort of psychological connectedness and/or continuity invoked in his criterion of personal identity over time—and in premise (II)—to be substantive. So in what follows, let us take claims about psychological connectedness and/or continuity to be claims about substantive psychological connectedness and/or continuity.

Suppose that you will undergo psychological fission. That is, suppose that you will be (substantively) psychologically connected to and/or continuous with two future persons. Then you are not related to either of those two future persons by unbranching psychological connectedness and/or continuity. So you are not related to either of those two future persons by what Parfit takes to be the criterion of personal identity over time. So Parfit would say that you are not the same person as either of those two future persons. Thus Parfit (1971; 1984, 253–66) takes the possibility of psychological fission to show:

> (IV) It is possible for you to be psychologically connected to and/or continuous with a person at a future time even if that person is not the same person as you.

Premises (III) and (IV) imply:

> (V) It is possible for a person at a future time to have, at that time, what matters in survival for you even if that person is not the same person as you.

This is Parfit's argument for the claim that personal identity is not necessary for what matters in survival.

In Section I, I argued that personal identity is necessary for what matters in survival. So Section I's argument is a reason to conclude that Parfit's argument to the contrary must go wrong somewhere or other. But if I relied on this reason for concluding that Parfit's argument goes wrong, you might tollens where I ponens. That is, you might take Parfit's argument to be a reason to deny that personal identity is necessary for what matters in survival, and so a reason to conclude that Section's I's argument must go wrong somewhere or other. I do not want to leave us at this impasse. So Sections III and IV will show right where Parfit's argument goes wrong. That is, they will show right where Parfit's argument goes wrong, given this or that metaphysics of persistence, and given this or that view about the criterion of personal identity over time.

III. Parfit's Argument, Stage Theory, and Perdurance

Suppose that stage theory is true. Then to enjoy personal identity with a person at a future time is to be I-related to that person at that time. That is, to be the same person as a future person is to be I-related to that person. As emphasized in Chapter 2 (§II), the I-relation is not numerical identity. Suppose that the I-relation is instead psychological connectedness and/or continuity. Then stage theorists will say that to be the same person as a future person is to be psychologically connected to and/or continuous with that person. So they will deny the following premise of Parfit's argument:

> (IV) It is possible for you to be psychologically connected to and/or continuous with a person at a future time even if that person is not the same person as you.

Again, suppose the I-relation is psychological connectedness and/or continuity. Then stage theorists should take the criterion of personal identity over time to be psychological connectedness and/or continuity. As we saw above (§II), Parfit endorses a different criterion of personal identity over time. Parfit's criterion is *unbranching* psychological connectedness and/or continuity. Parfit motivates his criterion with:

> In the imagined case where I divide, [psychological connectedness and/or continuity] takes a branching form. But personal identity cannot take a branching form. I and the two resulting people cannot be one and the same person. (1984, 262)

Stage theorists can resist Parfit's motivation for his criterion. This is because stage theorists can say that when psychological connectedness and/or continuity branches, so does personal identity.

Suppose that you will undergo psychological fission. And suppose that the I-relation is psychological connectedness and/or continuity. Then—according to stage theory—you will be the same person as each of the two stages, located at some future time, to which you are I-related. So you are the same person as Stage 1 and you are the same person as Stage 2. Stage 1 is not numerically identical with Stage 2. But since the I-relation is not numerical identity, but is instead psychological connectedness and/or continuity, none of this implies a contradiction.

Not only can stage theorists say that you will be same person as each person who results from psychological fission, this is just what the leading stage theorists—Theodore Sider (1996; 2001, 201) and Katherine Hawley—actually do say. For example, Hawley says:

> . . . there are puzzle cases which encourage us to think that people could divide . . . this is most easily accepted by those who believe that personal persistence is a matter of psychological continuity . . . stage theorists have a relatively simple strategy: there is but a single person present before fission, although it may be that she is the same person as two distinct but simultaneous person-stages at some post-fission moment. (2001, 174–5)

Let me illustrate the stage theorist's view here with one of Parfit's (1971) most important thought experiments (which he borrows from Wiggins (1967)). Suppose that each hemisphere of your brain is transplanted into a different body. Parfit insists that you would survive as each of the hemisphere recipients. Stage theorists can agree. Parfit thinks this implies that personal identity is not necessary for survival. Hawley and Sider disagree. They say that you survive as—and are the same person as—each recipient of each hemisphere. For you are I-related to each of those recipients.

The leading stage theorists claim that personal identity takes a branching form in cases of psychological fission. This claim undermines Parfit's reason for taking the criterion of personal identity over time to include an 'unbranching' condition. More to the point, this claim allows stage theorists to deny premise (IV) of Parfit's argument. So this claim allows stage theorists to block Parfit's argument.

Stage theorists say that your being the same person as a future person just is your being I-related to that person. Suppose the I-relation is a psychological relation. Then stage theorists say that your being the same person as a future person just is your being related to that person by way of the relevant psychological relation. So stage theory fits perfectly with Parfit's (e.g. 1984, 262–3) view that a future person's being the same person as you 'just is' your being related to that person by way of the relevant psychological relation. Moreover, stage theory is the only metaphysics of persistence that fits perfectly with that view. (Chapter 2 (§III) showed that endurance implies that that view is false; for perdurance and that view, see below.) So if we take Parfit completely literally when he says that a future person's

being the same person as you 'just is' your being related to that person by way of the relevant psychological relation, we should interpret Parfit as a (proto-) stage theorist. So Parfit should find the ease with which stage theorists can block his argument to be particularly troubling.

Suppose that persons perdure. And suppose that perduring persons enjoy a criterion of personal identity over time in terms of psychological connectedness and/or continuity. Then psychological connectedness and/or continuity is the 'glue' that holds temporal parts together so that those parts compose a single perduring person. Thus—as noted in Chapter 2 (§III)—perdurantists such as Lewis (1976) say that the relata of psychological connectedness and/or continuity are the temporal parts of persons, as opposed to perduring persons themselves.

And this is what all perdurantists should say. For I have argued that perdurantists should say that only the temporal parts of persons (and not perduring persons themselves) have (determinate) mental properties (Ch. 2, §II). Because temporal parts alone have (determinate) mental properties, temporal parts alone can be psychologically connected to and/or continuous with each other. So if persons perdure, the relata of psychological connectedness and/or continuity are the temporal parts of persons, as opposed to perduring persons themselves.

Perdurantists should take psychological connectedness and/or continuity—unbranching or otherwise—to hold not between persons, but instead between the temporal parts of persons. So perdurantists should deny that you—a person—can be psychologically connected to and/or continuous with a person at a future time. This is why perdurance—unlike stage theory—does not fit perfectly with Parfit's view that your being the same person as a person at a future time 'just is' your being related to that person by way of the relevant psychological relation.

Again, perdurantists should deny that you can be psychologically connected to and/or continuous with a person at a future time. So perdurance alone seems to rule out the claim that it is possible for you to be psychologically connected to and/or continuous with a person at a future time. A fortiori perdurance alone seems to rule out:

(IV) It is possible for you to be psychologically connected to and/or continuous with a person at a future time even if that person is not the same person as you.

But consider the following:

> (4D-IV): It is possible for your current temporal part to be psychologically connected to and/or continuous with a temporal part at a future time, but for it to be false that those two temporal parts are parts of the same person.

Perdurance alone does not rule out (4D-IV). Moreover, perdurantists might take (IV) to be shorthand for (4D-IV) (cf. Ch. 2, §III). Then perdurance alone does not rule out (IV) after all, at least not if (IV) is understood as (4D-IV).

Suppose that (IV) is understood as (4D-IV). And suppose that psychological fission is possible. Perdurantists can still deny (IV). One way they can deny (IV) is by claiming that psychological fission involves two distinct perduring persons who overlap before fission (see, e.g. Lewis, 1976). That is, those two persons share all their pre-fission temporal parts. This renders cases of psychological fission consistent with the claim that, necessarily, if any temporal part is psychologically connected to and/or continuous with another temporal part, then there is a person who has those two temporal parts as parts. So this renders cases of fission consistent with the denial of (4D-IV). So this renders cases of fission consistent with the denial of (IV) as understood by perdurantists.[5]

Let me illustrate this with Parfit's double brain hemisphere transplant case. Perdurantists can claim that your current temporal part is a part of the perduring person who is the recipient of your left hemisphere. They can also consistently add that your current temporal part is a part of the distinct perduring person who is the recipient of your right hemisphere. Your current temporal is, therefore, a temporal part of two distinct perduring persons. There is a sense, then, in which personal identity—at least from the perspective of your current temporal part—can take a branching form.

[5] Here is another option: there is a single perduring person in a case of psychological fission, who has a branching structure. This option also renders cases of fission consistent with the claim that, necessarily, if any temporal part is psychologically connected to and/or continuous with another temporal part, then there is a person who has those two temporal parts as parts.

Perdurantists can deny (IV). That is, perdurantists can deny (IV) even given the possibility of psychological fission. And they can deny (IV) even if they interpret (IV) as the perdurantist-friendly (4D-IV). So perdurantists can join stage theorists in concluding that Parfit's argument goes wrong by having a false premise, namely, (IV).

IV. Parfit's Argument and Endurance

Parfit's argument needs this premise:

> (IV) It is possible for you to be psychologically connected to and/or continuous with a person at a future time even if that person is not the same person as you.

Parfit's (1984, 262) support for premise (IV) assumes that 'personal identity cannot take a branching form'. And if persons persist by enduring, then personal identity cannot take a branching form.

To see why I say this, suppose—for reductio—that personal identity branches. In particular, suppose that you persist as Person 1 and also persist as the numerically distinct Person 2. Add that you persist by enduring. Then you are numerically identical with Person 1. And you are numerically identical with Person 2. So, by the transitivity of identity, Person 1 is numerically identical with Person 2. But Person 1 is not numerically identical with Person 2. Contradiction.

Endurantists cannot block Parfit's argument by claiming that personal identity branches in cases of psychological fission. But they can still block that argument. I block that argument by denying its first premise. That is, I deny:

> (I) If you are psychologically connected to and/or continuous with a person—but only that person—at a future time, then that person has, at that time, what matters in survival for you.

My reason for denying (I) starts with the claim that you have physical properties. In particular, I think that you have exactly the physical properties that are had by the human organism 'associated' with you. So you and that human organism have exactly the same physical parts. So you and that organism are in the same place at the same time. Add that an entity's mental properties are correlated with that entity's

physical properties and physical parts. Then you and that organism have the same mental properties.

You have the same physical parts and the same physical properties as the organism associated with you. So I conclude that you are not only a person, but are also an organism. The organism associated with you has the same mental properties as do you. So I conclude that that organism is not only an organism, but is also a person. I do not think that you and the organism associated with you are an instance of two organisms and two persons being in the same place at the same time. So I conclude that you are numerically identical with the human organism associated with you. (For more, see Merricks, 2001a, 47–53 and 85–7; Olson, 1997, 97–102; Olson, 2003).

Each of us is a human organism. Human organisms can enjoy identity over time in the absence of psychological connectedness and continuity. For example, a human organism can persist through the sort of brain damage that would keep the way it is now from being psychologically connected to, or continuous with, the way it will be after that damage. So it is false that unbranching psychological connectedness and/or continuity is necessary and sufficient for the identity over time of us human organisms. So it is false that unbranching psychological connectedness and/or continuity is the criterion of personal identity over time. Parfit's motivation for premise (I) relies on this false claim (see §II). So I conclude that Parfit does not successfully motivate premise (I). This alone is enough to block's Parfit's argument.[6]

Moreover, I think that premise (I) is false. For I deny that you could move a cat from Scotland to my pocket by annihilating that cat's body and then 'downloading' its psychological features onto my phone. So I deny that unbranching psychological connectedness and/or continuity is sufficient for the persistence of an organism, even an organism

[6] The argument in the text relies on the claim that you have the same physical properties as are had by the human organism associated with you. But suppose you think, instead, that you have the same physical properties as are had by some other physical object associated with you (e.g. your brain). Then conclude that you are numerically identical with that other physical object (e.g. that brain). I deny that that physical object (e.g. that brain) enjoys a psychological criterion of identity over time.

that has mental properties. Again, each of us is an organism. So I conclude that unbranching psychological connectedness and/or continuity is not sufficient for personal identity over time. So it is not sufficient for what—as I argued in Section I—is necessary for what matters in survival. So it is not sufficient for what matters in survival. So I deny:

> (I) If you are psychologically connected to and/or continuous with a person—but only that person—at a future time, then that person has, at that time, what matters in survival for you.[7]

In Section III, I discussed how stage theorists and perdurantists would block Parfit's argument. I did not suggest that stage theorists and perdurantists might deny premise (I). This is because I think that stage theorists and perdurantists are unlikely to deny premise (I). They are unlikely to deny (I) because perdurance undermines my argument for our being organisms. And stage theory provides a way to render our being organisms consistent with a psychological criterion of personal identity over time.

[7] You might think that if you are a human organism, then the criterion of personal identity over time is in terms of biological continuity. You might add that biological continuity can 'branch'. So you might think that if you are an organism, then the criterion of personal identity over time is unbranching biological continuity. You might then conclude that Parfit-style reasoning delivers the conclusion that biological continuity is sufficient for what matters in survival, and so personal identity is not necessary for what matters in survival.

As we shall see below, endurantists can block Parfit's argument even if they accept Parfit's criterion of personal identity over time in terms of unbranching psychological connectedness and/or continuity. This will show how endurantists can block Parfit-style reasoning against the necessity of personal identity for survival even if they accept that the criterion of personal identity over time is unbranching biological continuity. That is the main point.

But other points are also relevant. I myself deny that there is any criterion of personal identity over time at all (Merricks, 1998 and 2001b, 191–5); so I deny that the criterion of personal identity over time is in terms of biological continuity, unbranching or otherwise. Moreover, it is not at all clear that biological continuity of the sort that is even prima facie sufficient for the persistence of an organism can branch. (A dividing cell does not persist.) In particular, biological continuity does not branch in Parfit's double brain hemisphere transplant case, since the brain stem and the original human organism get left behind.

Suppose that objects persist by perduring. And suppose that the perduring human organism that is associated with you has a proper part that itself perdures, and all of whose temporal parts are related by psychological continuity. Perdurantists can claim that you are (identical with) this proper part of the organism associated with you. So you do not have all the same physical properties and parts as that organism. (For example, that organism, but not you, has early fetal temporal parts.) So perdurantists can deny that you are an organism. More generally, they can deny that human persons are organisms (for more, see Merricks, 2000; Olson, 1997, 162–8).

Suppose that stage theory is true. Let the I-relation be psychological connectedness and/or continuity. Let the O-relation be biological continuity. Stage theorists say that your being the same person as a future stage just is your being I-related to that future stage. They can add that your being the same organism as a future stage just is your being O-related to that stage. The O-relation is not the same relation as the I-relation. So stage theorists can say that you are I-related to one stage that is located at a later time, but O-related to a numerically distinct stage located at that same time. So stage theorists can say that you persist *qua* person as one stage, but persist *qua* organism as another. So stage theorists can—and do—take persistence to be sortal-relative (see Sider, 1996; Hawley, 2001, 156–8).

So stage theorists can consistently endorse the following four claims. Each of us is the same organism as some future stage; so each of us is an organism. The criterion for being the same organism as some future organism is not psychological. Each of us is the same person as some future stage; so each of us is a person. And the criterion for being the same person as some future person is psychological. Because stage theorists can consistently endorse these four claims, they can deny that each of us being an organism implies that the criterion of personal identity over time is not psychological.

Suppose that persons endure. Then I think that my above argument for the claim that each of us is a human organism succeeds. This is partly because if persons endure, it is false that each person is a perduring proper part of a perduring human organism. Again, suppose that persons endure. Then I think that our being organisms implies that it is false that the criterion of our identity over time—and so the criterion of personal identity over time—is psychological. This is

partly because if persons endure, then it is false that persistence is sortal-relative. This is false because numerical identity is not sortal-relative.[8]

At any rate, I think that endurantists in particular should deny that the criterion of personal identity is psychological. But maybe you disagree. So let us assume, just for the sake of argument, that persons endure but the criterion of personal identity over time is psychological. Indeed, let us even concede to Parfit that the criterion of personal identity over time is the very criterion that his argument requires: unbranching psychological connectedness and/or continuity.

Recall from Section II that Parfit's motivation for (I) turns on two claims. The first is that the criterion of personal identity over time is unbranching psychological connectedness and/or continuity. The second is that personal identity over time is sufficient for what matters in survival. As we saw in Chapter 2 (§I), endurantists in particular should endorse this second claim. So—assuming Parfit's criterion of personal identity over time for the sake of argument—endurantists should endorse:

> (I) If you are psychologically connected to and/or continuous with a person—but only that person—at a future time, then that person has, at that time, what matters in survival for you.[9]

But endurantists who accept (I) should deny:

> (II) If premise (I) is true, then: if you are psychologically connected to and/or continuous with a person at a future time, then that person has, at that time, what matters in survival for you.

[8] Geach's (1980, 216) 'relative identity' might seem to be a view according to which numerical identity is sortal-relative. But Geach's view is really the view that there is no such thing as numerical identity, but instead a variety of sortal-relative identity-like relations. For critical discussion of 'relative identity', see Hawthorne (2003).

[9] Endurantists who accept (I) should take (I) to be shorthand for something like: if the way you are now is psychologically connected to and/or continuous with the way a (conscious) person is at a future time—and only with the way that that person is at that time—then that person has, at that time, what matters in survival for you. Something similar goes for premises (II), (III), and (IV). (See Ch. 1, §III and Ch. 2, §III.)

To begin to see why I say that endurantists who accept (I) should deny (II), consider how Parfit motivates (II):

> If I will be [psychologically connected to and/or continuous with] some future person, [whether that continuity and/or connectedness] branches makes no difference to the intrinsic nature of my relation to this person. And what matters most must be the intrinsic nature of this relation. (1984, 263)

In one scenario, you are related to a person at a future time by way of unbranching psychological connectedness and/or continuity. In a second scenario, you are related to that person at that time by way of one branch of branching psychological connectedness and/or continuity. Parfit assumes that these scenarios differ in how you are related to that person at that future time only with regard to whether branching occurs, and so only 'extrinsically'. And Parfit's assumption would be right if stage theory were true.

But if persons endure, Parfit's assumption is wrong, at least assuming that Parfit's criterion of personal identity over time is right. For assuming his criterion, in the first scenario you are related to that person at that future time by way of *being numerically identical with*, and in the second you are not. And this is a crucial difference. For, as we saw in Chapter 2 (§I), endurantists should say that your being numerically identical with a person at a future time explains why that person will have, at that future time, what matters in survival for you.

Again, endurantists say that your being numerically identical with a person at a future time explains why that person will have, at that future time, what matters in survival for you. They do not thereby say that this is explained by your being related to that person by the criterion of personal identity over time. So even those endurantists who accept Parfit's criterion are not thereby saying that survival is explained by being related by unbranching psychological connectedness and/or continuity. Perhaps being related by Parfit's criterion might also explain survival. But being related by that criterion would explain survival—if at all—only by explaining being related by the 'further fact' of numerical identity, which is what directly explains survival (see Ch. 2, §III).

So endurantists who accept Parfit's criterion of personal identity over time have no reason to accept:

(II) If premise (I) is true, then: if you are psychologically connected to and/or continuous with a person at a future time, then that person has, at that time, what matters in survival for you.

So endurantists who accept Parfit's criterion of personal identity over time can deny (II) at no cost. So there is nothing preventing them from denying (II) if they are motivated to do so. And endurantists who believe—perhaps because of the argument of Section I—that numerical identity is necessary for survival are motivated to deny (II). For given Parfit's criterion of personal identity over time, it is false that being psychologically connected to and/or continuous with a person is sufficient for personal identity over time. So it is false that it is sufficient for numerical identity, which is necessary for survival. So it is false that it is sufficient for survival.

You might object that endurantists do have a reason to accept both (I) and also (II). For you might object that endurantists should take the criterion of personal identity over time to be psychological connectedness and/or continuity, leaving behind Parfit's unbranching clause. And you might even think that Parfit should leave that clause behind, since you might think that Parfit should say that what explains the holding of personal identity over time ought to be 'intrinsic'. But endurantists cannot take the criterion of personal identity over time to be psychological connectedness and/or continuity. For we are assuming—as Parfit must—that psychological connectedness and/or continuity can branch. So the claim that the criterion of personal identity over time is psychological connectedness and/or continuity implies—given endurance—that numerical identity can branch. But numerical identity cannot branch, on pain of contradiction.

Let us apply all this to Parfit's double brain hemisphere transplant case. You will undergo double brain hemisphere transplant. So psychological connectedness and/or continuity will branch. So it is false that either hemisphere will be related to you by way of unbranching psychological connectedness and/or continuity. Will you survive this procedure?

It is not yet clear. For—given endurance—the case has not been fully described. We have not yet said to which, if either, of the resulting persons you will be related by the 'further fact' of numerical

identity. Nevertheless, we do know that you cannot be related by numerical identity to both (numerically distinct) hemisphere recipients. So endurantists convinced by Section I's argument for the necessity of identity for survival should deny that you survive as both hemisphere recipients.

Suppose we add—just for the sake of argument—that unbranching psychological connectedness and/or continuity is the criterion of personal identity over time. Then unbranching psychological connectedness and/or continuity is necessary for numerical identity. Then we should conclude that you are related to neither hemisphere recipient by the 'further fact' of numerical identity. Then endurantists convinced by Section I's argument for the necessity of identity for survival should deny that you survive as either hemisphere recipient.

Or suppose we add—again, for the sake of argument—that you are a human organism. Suppose that each recipient of one of your brain hemispheres will also be a human organism. But neither recipient will be the same human organism as you. For removing your brain hemispheres still leaves behind the brainstem, and so seems to leave behind the original human organism. So we should conclude that you are related to neither hemisphere recipient by the 'further fact' of numerical identity. Again, endurantists convinced by Section I's argument for the necessity of identity for survival should deny that you survive as either hemisphere recipient.

Parfit's argument starts with these two premises:

> (I) If you are psychologically connected to and/or continuous with a person—but only that person—at a future time, then that person has, at that time, what matters in survival for you.
>
> (II) If premise (I) is true, then: if you are psychologically connected to and/or continuous with a person at a future time, then that person has, at that time, what matters in survival for you.

Endurantists can join me in denying (I). And those who do not join me—but who instead endorse (I)—have no reason at all to endorse (II), and so can deny (II). So all endurantists can block Parfit's argument for the conclusion that personal identity is not necessary for survival.

This section and Section III have shown that everyone—regardless of one's metaphysics of persistence—can block Parfit's argument.

These sections have also shown that the particular way in which one will block that argument depends on one's metaphysics of persistence, as well as one's view on the criterion of personal identity over time. But the main point is that everyone can block that argument. So that argument fails to establish its conclusion. So that argument fails.

Of course, sometimes even a failed argument happens to have a true conclusion. So neither this section nor Section III shows that the conclusion of Parfit's argument is false. Relatedly, neither this section nor Section III shows that it is false that psychological connectedness and/or continuity is sufficient for survival, regardless of whether personal identity holds. But do not forget Section I. In Section I, I argued for the claim that personal identity is necessary for what matters in survival. This claim implies that Parfit's conclusion that personal identity is not necessary for survival is false. And this claim implies that it is false that psychological connectedness and/or continuity is sufficient for survival, regardless of whether personal identity holds.

V. Psychological Connectedness and Psychological Continuity

I do not have more to say in defense of the claim that personal identity is necessary for what matters in survival. So I do not have more to say in defense of the claim that if persons endure, then every good answer to the following implies being numerically identical with:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

So Chapters 4–6 will not further defend the claim that every good answer to the Why Question implies numerical identity. Rather, those chapters will further defend the claim that being numerically identical with is a good answer to the Why Question.

As Chapter 2 (§III) pointed out, many assume that an answer to the Why Question should be in terms of the criterion of personal identity over time, and that that criterion should therefore be in terms of psychological connectedness or psychological continuity. This implies that many assume that an answer to the Why Question should be in terms of psychological connectedness or psychological continuity.

Let us say that an answer to the Why Question 'competes' with my answer just in case if that competing answer were a good answer, then it would be false that numerical identity is a good answer to the Why Question. I take the most serious competing answers to be in terms of—or at least imply—psychological connectedness or psychological continuity. This is because I think that almost all philosophers who do not endorse my answer to the Why Question will instead answer the Why Question in terms of—or at least in a way that implies—some sort of psychological connectedness or psychological continuity.

My goal in Chapters 4–6 is to defend my answer to the Why Question from the threat of competing answers to that question. Because this is my goal in those chapters, those chapters will not focus on what is most important about this or that philosopher's overall view, much less on how a philosopher's view has evolved over time and publication. Rather, those chapters will instead focus only on that aspect of this or that philosopher's work that delivers a competing answer to the Why Question, and in particular that delivers a competing answer in terms of—or that at least implies—psychological connectedness or psychological continuity.

One example of psychological connectedness is memory: the way you are now is psychologically connected to the way a person was at a past time if you now remember the experiences that that person had at that past time. Another example involves intentions to act: the way you are now is psychologically connected to the way a person will be at a future time if that person will, at that time, act on your current intention. Yet another example of psychological connectedness is being psychologically alike in some way, such as having the same values and desires. And psychological continuity is defined in terms of psychological connectedness. The way a person is at an earlier time being psychologically continuous with the way a person is at a later time just is the way a person is at an earlier time being related by a chain of overlapping instances of psychological connectedness to the way a person is at a later time.

Again, Chapters 4–6 focus on competing answers to the Why Question that are all in terms of—or imply—psychological connectedness or psychological continuity. But unlike Parfit's answer in terms of generic psychological connectedness and/or continuity, the answers considered in the following chapters are in terms of—or imply—this or that specific sort of psychological connectedness or psychological

continuity. These 'specific sorts' include having the 'same self' (Chs 4 and 5); having the 'same self-narrative' (Ch. 5); being continuous with regard to overlapping 'local narratives' (Ch. 6); and being continuous with regard to overlapping 'agential connectedness' (Ch. 6).

I shall show that each of these answers to the Why Question really does compete with my answer. So my answer competes with each of these answers. So I conclude that the argument in Chapter 2 (§I) for my answer to the Why Question is thereby an argument for the falsity of each of these competing answers. So I conclude that each of these competing answers is false.

Of course, you might tollens where I ponens. That is, you might endorse one or another of the answers to the Why Question considered in Chapters 4–6. Then you might take your answer's competing with my answer to be a reason to reject my answer, and so a reason to conclude that my argument for that answer must somehow go wrong.

I do not want to leave us at this impasse. So Chapters 4–6 will not only object to competing answers by way of my argument for my answer to the Why Question. Chapters 4–6 will also object to competing answers in ways that do not presuppose my answer to the Why Question, or my argument for that answer. Those objections cannot be blocked simply by rejecting what I say about the Why Question.

Chapters 4–6 are intended to rule out a number of competing answers to the Why Question that are in terms of—or that imply—specific sorts of psychological connectedness or psychological continuity. And I think that the problems with those competing answers will motivate a general moral. This is the moral that we should reject any answer to the Why Question that both competes with my answer and also is in terms of—or implies—psychological connectedness or psychological continuity, generic or otherwise. I defend this moral at the end of Chapter 6 (§VI).

VI. Conclusion

Suppose that persons endure. Then—as I have argued (§I)—being numerically identical with is implied by all good answers to:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

And so, if persons endure, personal identity is necessary for survival. Suppose that persons do not endure. Then there are still reasons for taking personal identity—that is, being the same person as—to be necessary for survival (§I). At any rate, I conclude that personal identity is necessary for survival. And we have seen that Parfit's argument to the contrary can be blocked by anyone who wants to block it (§§II–IV).

I answer the Why Question in terms of numerical identity. My defense of my answer is twofold. The first aspect of that defense is my positive argument for my answer, already given in Chapter 2 (§I). The second aspect of that defense is spread across the next three chapters. That second aspect is an extended argument for the conclusion that we should reject the most attractive alternatives to my answer to the Why Question, and reject them for reasons that have nothing to do with my answer.

4

The Same Self

Some philosophers say that a person will have, at a future time, what matters in survival for you only if the way you are now is relevantly psychologically connected to the way that person will be at that time. For example, Leibniz says this, taking the relevant sort of psychological connectedness to be memory (Ch. 1, §§I–II). This chapter focuses on others who say this but who, unlike Leibniz, take the relevant sort of psychological connectedness to be being substantively alike with regard to values, desires, and projects.

I. Three Selfers

John Perry says:

> Most of us have a special and intense interest in what will happen to us. You learn that someone will be run over by a truck tomorrow; you are saddened, feel pity, and think reflectively about the frailty of life; one bit of information is added, that the someone is you, and a whole new set of emotions rise in your breast. (1976, 67)

'Special and intense interest in what will happen to us' is Perry's way of describing future-directed self-interested concern.

Perry adds:

> I expect to have tomorrow much the same desires, goals, loves, hates—in a word, projects—as I have today. There is, in the normal case, no one as likely to work on my article, love my children, vote for my candidates, pay my bills, and honor my promises, as me. (1976, 75)

Self and Identity. Trenton Merricks, Oxford University Press. © Trenton Merricks 2022.
DOI: 10.1093/oso/9780192843432.003.0005

'In the normal case', you and only you will be likely to work on your current projects. And Perry thinks this is what justifies your having future-directed self-interested concern with regard to your own future experiences.

Consider what Perry takes to be an abnormal case. You will continue to exist but will change in ways that will make you no longer likely to work on your current projects. Then Perry (1976, 78–80) denies that there is a 'justification' for your having future-directed self-interested concern with regard to your own future experiences.

Suppose that you will cease to exist and be replaced by—in Perry's (1976, 79–80) words—a 'benign imposter'. That is, you will be replaced by someone who is not you but who will work on your current projects. Then Perry thinks that your having future-directed self-interested concern with regard to this imposter would be justified. Perry (1976, 80) even adds that your withholding such concern would be 'irrational'.

So Perry (1976) would say that it is appropriate for you now to have future-directed self-interested concern with regard to a person's experiences at a future time if and only if that person will be likely, at that time, to work on your current projects. Perry thinks that a person will be likely to work on your current projects at a future time if and only if that person will then have the 'same desires, goals, loves, hates' as you now have. So Perry would say that it is appropriate for you now to have future-directed self-interested concern with regard to a person's experiences at a future time if and only if that person will then have the 'same desires, goals, loves, hates' as you now have.

Marya Schechtman (1996, 2) takes your future-directed self-interested concern to be appropriate if and only if it is directed at a person at a future time who will have, at that time, the 'beliefs, values, desires and other psychological features [that] make [you] the person [you are]'. As we shall see in Chapter 5 (§I), Schechtman thinks that the beliefs, values, desires, and other psychological features that 'make you the person you are' are those that are included in your self-narrative. So Schechtman would say that it is appropriate for you now to have self-interested concern with regard to a person's experiences at a future time just in case that person will have, at that time, the beliefs, values, desires, and so on that are now included in your self-narrative.

In 'Friends and Future Selves', Jennifer Whiting says:

> I have special reasons to care about my future self if her experiences are connected in certain important ways to my present ones—if, for example, she will recall and carry out some of my present but future-directed intentions. (1986, 547)

Thus Whiting takes certain sorts of psychological connectedness to be sufficient for the appropriateness of future-directed self-interested concern. Whiting's example is acting on remembered intentions.

Acting on remembered intentions is correlated with—or might even just be—working on the same projects. For, at least typically, someone will continue to work on your article just in case that person will act on a remembered intention to work on that article; and someone will honor your current promises just in case that person will act on a remembered intention to keep the promises you are now making; and so on. Thus Whiting, like Perry, seems to take working on your current projects to be sufficient for the appropriateness of future-directed self-interested concern.

Moreover, Whiting seems to take having (enough of) the same values and desires to be necessary for the appropriateness of future-directed self-interested concern. For she says:

> The difficult cases are those where our . . . future selves have desires or interests which express values that conflict with those we now believe to be justified—for example, their desires to join the Ku Klux Klan . . . at some point we may simply say that their desires are so morally outrageous that our concern for [those future selves] is no longer justified. (1986, 559)

There are differences between Perry, Whiting, and Schechtman. For example, Perry focuses on the psychological features that are bound up with being likely to work on your projects, while Schechtman focuses on those that are included in your self-narrative. But do not let such differences obscure the significant way in which Perry, Whiting, and Schechtman are alike. Perry, Whiting, and Schechtman are each committed to the claim that your future-directed self-interested concern with regard to a person's experiences at a future time is appropriate if and only if the way you are now is psychologically connected to the way that that person will be at that time. And each of them takes the relevant psychological connectedness to be a matter of being

psychologically alike. In particular, each of them takes this to be a matter of being alike with regard to some combination or other of values, desires, and those psychological features that make one likely to work on—or constitute one's having—particular projects. So let us say that each of them takes the relevant psychological connectedness to be a matter of being alike with regard to 'values, desires, and projects'.

Perry, Whiting, and Schechtman say that your future-directed self-interested concern with regard to a person's experiences at a future time is appropriate if and only if that person will have, at that time, enough of your current values, desires, and projects. It is appropriate for you to have future-directed self-interested concern with regard to a good (or bad) experience if and only if it is appropriate for you to first-personally anticipate that experience (Ch. 1, §II). So Perry, Whiting, and Schechtman should all agree that a person at a future time will have, at that time, what matters in survival for you if and only if that person will have, at that time, enough of your current values, desires, and projects. They should all agree to this because of the answer to:

> *The What Question*: What is it for a person at a future time to have (at that time) what matters in survival for you?

The answer: its being appropriate for you to first-personally anticipate the experiences that that person will have at that future time; and if that person will have good (or bad) experiences at that time, its being appropriate for you to have future-directed self-interested concern with regard to those experiences (see Ch. 1, §II).

Perry does not say how likely a future person must be to work on your projects in order to constitute the relevant psychological connectedness. Nor does Perry say how many of your projects a future person must be likely to work on. And neither Schechtman nor Whiting tells us how many values and desires and projects are required to secure the relevant psychological connectedness. So none of Perry, Schechtman, or Whiting specifies the degree of psychological connectedness that they take to be necessary and sufficient for survival. But I think all of them would agree that the relevant psychological connectedness must be quite substantive.

To see why I say this, suppose that every person always has the desire to be happy. Then each person who exists at the present time has a desire that will also be had by each person who will exist at every

future time. Then each person at the present time is to some degree psychologically connected to each person at every future time. But surely Perry, Whiting, and Schechtman deny that each person at every future time will have, at that time, what matters in survival for you, and for me, and for every other person who now exists. So the psychological connectedness they have in mind must be more substantive than being alike with regard to a single desire.

I assume that Perry, Whiting, and Schechtman take the psychological connectedness that is necessary and sufficient for what matters in survival to be so substantive that—normally—each person will be thus psychologically connected to at most one person at a future time. For I assume that Perry, Whiting, and Schechtman would say that—normally—at most one person at a future time will have what matters in survival for you.

One or another of them might think that after each hemisphere of your brain is transplanted into a new body, then two persons will have what matters in survival for you; but this surgical procedure is abnormal. Perry might think that if you will continue to exist and pursue your current projects alongside a benign imposter, then two persons will have what matters in survival for you; but benign imposters are abnormal.

Let us stipulate that a person will have, at a future time, the *same self* as you now have just in case the way you are now is—with regard to values, desires, and projects—substantively psychologically connected to the way that person will be at that future time. How substantive? So substantive that—abnormal cases aside—at most one person at a future time will have the same self as you now have. Note that having the same self—given the above stipulation—does not imply that there is any such entity as a 'self'. Rather, to have the same self just is to be substantively alike with regard to values, desires, and projects.

Let us also stipulate that you will undergo a *change in self* if you persist until a future time but will not have, at that time, the same self as you have now. In other words, you will undergo a change in self if the way you are now is not—with regard to values, desires, and projects—substantively like the way you will be at a future time.

Finally, let us stipulate that to be a *Selfer* is to be committed to the view that a person at a future time will have, at that time, what matters in survival for you only if that person will be substantively like you with

regard to values, desires, and projects. That is, to be a *Selfer* is to be committed to the view that a person at a future time will have, at that time, what matters in survival for you only if that person will have, at that time, the same self as you now have. So Selfers are committed to the view that a person cannot survive a change in self.

This chapter focuses on the Selfer view. That is, this chapter focuses on the view that having the same self is necessary for survival. This is because I think that the Selfer view is even more widely endorsed than is the stronger claim that having the same self is both necessary and sufficient for survival. But we have seen that Perry, Whiting, and Schechtman are committed to the claim that having the same self is both necessary and sufficient for survival. So Perry, Whiting, and Schechtman are Selfers, and then some.

Let us apply the Selfer view to an example from L. A. Paul's *Transformative Experience*:

> Imagine you have the chance to become a vampire. With one swift, painless bite, you'll be permanently transformed into an elegant and fabulous creature of the night. As a member of the undead, your life will be completely different. You'll experience a range of intense, revelatory new sense experiences, you'll gain immortal strength, speed, and power, and you'll look fantastic in everything you wear. You'll also need to drink blood and avoid sunlight.
>
> Suppose that all of your friends, people whose interests, views, and lives were similar to yours, have already decided to become vampires. And all of them tell you that they *love* it. They describe their new lives with unbridled enthusiasm, and encourage you to become a vampire too. They assuage your fears and explain that modern vampires don't kill humans; they drink the blood of cows and chickens. They say things like: 'I'd never go back, even if I could. Life has meaning and a sense of purpose now that it never had when I was human. I understand Reality in a way I just couldn't before.' (Paul, 2014, 1)

If you were to become a vampire of the sort just described, your life would acquire a new meaning and a new sense of purpose. I think that this implies that if you were to become this sort of vampire, you would not be likely to work on your current projects, or at least not enough of them to make the way you are now substantively psychologically connected to the way you will be as a vampire. Maybe you disagree. Then add to the above description that if you were to become a vampire you would not be likely to work on (enough of) your current projects. Perry (1976) thinks that a person will have, at a future time,

what matters in survival for you just in case that person will be likely, at that time, to work on (enough of) your projects. So Perry must say that you would not survive becoming a vampire.

I think that becoming a vampire is a paradigm case of changing with regard to what Schechtman (1996, 2) calls the 'beliefs, values, desires and other psychological features [that] make someone the person she is'. Maybe you disagree. Then add whatever it takes to the above description to make becoming a vampire a paradigm case of such a change. Schechtman (1996) thinks a future person will have what matters in survival for you just in case that person will have those features that, according to Schechtman, 'make you the person you are'. So Schechtman must say that you would not survive becoming a vampire.

Whiting thinks that a person will have, at a future time, what matters in survival for you just in case the way you are now is relevantly psychologically connected to the way that person will be at that future time. Whiting's (1986) examples of such connectedness are your having the same values and desires as that future person, and that future person's acting on your current intentions. Becoming the above sort of vampire is a paradigm case of changing in values and desires. And becoming this sort of vampire implies that you would not work on (enough of) your current projects, and so you would not act on (enough of) your pre-vampire intentions. So Whiting must say that you would not survive becoming a vampire.

The above Selfers must say that you would not survive becoming the sort of vampire described above. That is, these Selfers must say that no such vampire at a future time would have, at that time, what matters in survival for you. And I think that all Selfers must say this. For becoming the sort of vampire described above would result in your no longer being substantively like the way you are now with regard to values, desires, and projects. That is, to become this sort of vampire is to undergo a change in self.

The Selfer view implies that you would not survive becoming a vampire. I think that Selfers will be pleased with this implication. Again, the Selfer view has the result that even if you will be 'metaphysically' identical with a future vampire, that vampire will not really 'be you' in the way that matters. That is, that vampire will not have what matters in survival for you. I think that Selfers will take this result to be a benefit of their view.

At least typically, it is rational to choose to survive rather than to choose not to survive. (Paul (2014, 54) seems to agree.) Let us suppose that your current life is thus typical. That is, let us suppose that it is rational for you to choose to survive rather than to choose not to survive. The Selfer view implies that you would not survive becoming a vampire. So if you are given the option to become a vampire, the rational choice—by the Selfer's lights—is to decline.

The Selfer view is the view that you cannot survive a change in self. Add that it is rational for you to choose to survive rather than to choose not to survive. Then the Selfer view implies that whenever you are given the option to undergo a change in self, the rational choice is always to decline. So the Selfer view implies that whenever you are given the option to radically transform, the rational choice is always to decline.

Edna Ullmann-Margalit (2006) and L. A. Paul (2014, 2–3) have argued that there is no rational way to decide whether or not to undergo a radical transformation. Indeed, Paul's main reason for asking you to 'imagine that you have the chance to become a vampire' is to present you with a choice that—so Paul claims—cannot be made rationally. But we have just seen that the Selfer view implies otherwise. This shows that the Selfer view is substantive and controversial.

II. First-Personal Access to a Point of View

David Velleman (1996, 43) says: 'Our desire for a future to anticipate, I shall argue, is a desire for first-personal access to a future point of view.' I assume that if you will have what Velleman calls 'a future to anticipate', then it is appropriate for you to first-personally anticipate some future experiences. So I take Velleman to claim that it is appropriate for you to first-personally anticipate some future experiences just in case you have first-personal access to a future point of view. So I conclude that Velleman would claim that first-personally anticipating the experiences that a person will have at a future time is appropriate just in case you have first-personal access to that person's point of view at that time.

Appropriate first-personal anticipation goes hand in hand with appropriate future-directed self-interested concern (Ch. 1, §II). That is, it goes hand in hand with what Velleman (1996, 42) calls appropriate 'self-regarding concern about the future'. So Velleman should say

that 'self-regarding concern' about a person's experiences at a future time is appropriate just in case you have first-personal access to that person's point of view at that time. And he does seem to say this:

> My aim is to argue for this reinterpretation of our self-regarding concern about the future. What matters most, I shall suggest, is . . . whether there will be a future person whom I can now regard as self. And whether I can regard a future person as self, I shall argue, doesn't necessarily depend on whether he will be the same person as me; it depends instead on my access to his point of view. (1996, 42)

Velleman takes it to be appropriate for you to first-personally anticipate, and have future-directed self-interested concern with regard to, the experiences a person will have at a future time just in case you have first-personal access to that person's point of view at that time. So Velleman is committed to the claim that a person who will exist at a future time will have, at that time, what matters in survival for you just in case you now have first-personal access to that person's point of view at that future time. So it is no surprise that Velleman says:

> The future 'me' whose existence matters here is picked out precisely by his owning a point of view into which I am attempting to project my representations of the future, just as a past 'me' can be picked out by his having owned the point of view from which I have recovered representations of the past. (1996, 68)

Velleman thinks that you will survive as a person at a future time just in case you now have first-personal access to that person's point of view at that time. Velleman (1996) does not think that you have access to a person's point of view at a future time because you will be (identical with) the person who will have that point of view. Instead, Velleman thinks that you have first-personal access to a person's point of view in virtue of 'those psychological connections that function like memory in giving us first-personal access to other points of view' (1996, 42n5). So Velleman thinks that you now have first-personal access to a person's point of view at a future time just in case the way you are now is relevantly psychologically connected to the way that person will be at that future time.[1]

[1] Velleman (1996, 69–70) also claims that you have first-personal access to a point of view just in case you can 'unselfconsciously project' yourself into that point of view. I think that Velleman (1996) would say that you can unselfconsciously

Again, Velleman thinks that you will survive as a person at a future time just in case you now have first-personal access to that person's point of view at that time. We have just seen that Velleman thinks that you have first-personal access to a person's point of view at a future time just in case the way you are now is relevantly psychologically connected to the way that person will be at that future time. So Velleman must say that you will survive as a person at a future time only if you are psychologically connected to that person at that time. Selfers say this. Indeed, Velleman himself is (nearly) a Selfer. Or so I shall now argue.

Velleman says that memory gives you first-personal access to a past point of view. Velleman's main example of psychological connectedness that gives you first-personal access to a future point of view is framing an intention to act. Here is how Velleman characterizes framing an intention to act:

> Framing an intention entails projecting myself into a future perspective because it entails representing the intended action from the point of view of the agent who is to perform it. (1996, 70)

And:

> I attempt to frame the intention, if you will, from the intention's own future perspective, the perspective in which the intention itself will turn up to be executed. (1996, 71)

I cannot know what it is like to be a bat. So I cannot represent an intended action from the point of view of a bat. Nor can I frame an intention from 'the perspective in which it will turn up to be executed' by a future bat. So Velleman must say that I cannot frame an intention to act as a future bat. This illustrates a more general point. That point is that—given how Velleman characterizes framing an intention to act—if you cannot know what it would be like to be an entity at a future time, then you cannot frame an intention to act as that entity at that future time. This is true even if that entity is not a bat, but is instead a person.

project yourself into a person's point of view at a future time in virtue of the way you are now being relevantly psychologically connected to the way that person will be at that future time.

I do not know whether Velleman thinks that your having first-personal access to a future point of view implies that you actually frame intentions to act from that point of view. But he does seem to think that if you cannot frame an intention to act from a future point of view, then you do not have first-personal access to that point of view. So—given how he characterizes framing an intention to act—Velleman should say that if you cannot know what it would be like to be a person at a future time, then you do not have first-personal access to that person's point of view at that time.

Velleman's main example of psychological connectedness that gives you first-personal access to a future point of view is framing an intention to act. But Velleman (1996) discusses one other example: first-personally anticipating future experiences. So suppose that first-personally anticipating future experiences delivers first-personal access to the point of view that includes those experiences. Then—according to Velleman—it would deliver survival. Then first-personally anticipating an experience would make it appropriate to first-personally anticipate that experience. And this seems to be what Velleman says:

> And to imagine a pain as experienced by a mind hereby so prepared for it is already to brace for the pain, to shrink from it, or to be otherwise caught up in it in some way. Anticipation that's cognizant of its effect on the prefigured experiences is thus a form of mental engagement with them that, to some degree, already constitutes their mattering. (1996, 73)

Velleman seems to say that first-personally anticipating an experience makes it appropriate to first-personally anticipate that experience. This implies that first-personally anticipating an experience is always appropriate. One might object to Velleman here by claiming that first-personally anticipating an experience can fail to be appropriate. (Recall deluded Brown (Ch. 1, §I).) But I shall not pursue this objection. I shall instead continue to develop my argument for Velleman's being (nearly) a Selfer.

I cannot know what it is like to be a bat. So I cannot imagine a pain 'as experienced by' a bat's mind. That is, I cannot be 'cognizant' of a pain's 'effect on the prefigured experiences' of a bat. So I take Velleman's characterization of first-personal anticipation to imply that I cannot first-personally anticipate a bat's future pain. More importantly, Velleman's characterization implies that if you cannot

know what it would be like to be a person who is having an experience at a future time, then you cannot first-personally anticipate that person's experience at that time.

This reinforces the conclusion that Velleman must say that if you cannot know what it would be like to be a person at a future time, then you do not have first-personal access to that person's point of view at that time. I have belabored this conclusion because it is important. But it probably did not need belaboring. After all, anything that would plausibly count as 'having first-personal access to a point of view' seems to include being able to know what it would be like to have that point of view. I doubt that anyone would say: 'I cannot know what it is like to be a bat, but I do enjoy first-personal access to a bat's point of view.'

At any rate, Velleman must say that if you cannot know what it would be like to be a person at a future time, then you do not have first-personal access to that person's point of view at that future time. And we have seen that Velleman insists that if you do not have first-personal access to a person's point of view at a future time, then that person will not have, at that time, what matters in survival for you. So Velleman is committed to the claim that if you cannot know what it would be like to be a person at a future time, then that person will not have, at that time, what matters in survival for you. In other words, you will survive as a person at a future time only if you can know what it would be like to be that person at that time.

With all this in mind, let us return to Paul's vampires:

> [The vampires] say things like: 'I'd never go back, even if I could. Life has meaning and a sense of purpose now that it never had when I was human. I understand Reality in a way I just couldn't before. It's amazing. But I can't really explain it to you, a mere human—you have to be a vampire to know what it's like.' (Paul, 2014, 1)

You are not a vampire. So you cannot know what it would be like to be a vampire. So Velleman must deny that you would survive becoming a vampire.

Here is one reason that I think you cannot know what it would be like to be a vampire. You cannot know what it would be like to have values, desires, and projects that differ significantly from your values, desires, and projects. So you cannot know what it would be like to have the values, desires, and projects of a vampire. That is, you cannot know

what it would be like to be a vampire because you cannot know what it would be like to have a different self, or at least not a significantly different self. (Paul (2014; 2015a) goes so far as to claim that an adult woman who is not yet a mother cannot know what it would be like for her to have the values, desires, and projects she will have as a mother.)

Suppose that you are now a vampire. Back when you were a normal human, you could not know what it would be like to have your current vampiric values, desires, and projects. But you—a vampire—remember what it was like to have the values, desires, and projects of a normal human. So you know what it would be like to have those values, desires, and projects. So I need to qualify my remark that you cannot know what it would be like to have a different self. Let me say instead that you cannot know what it would be like to have a different self unless you had that self in the past and remember what it was like.[2]

You have first-personal access to a person's point of view at a future time only if you can know what it would be like to be that person at that future time. You cannot know what it would be like to have a different self (unless you had that self in the past). So if a person at a future time will have a different self than you have now (or have ever had), then you do not have first-personal access to that person's point of view at that future time. So—according to Velleman—that person will not have, at that future time, what matters in survival for you. Thus Velleman's views imply that you cannot survive a change in self (unless that change results in your reacquiring values, desires, and projects that you once had). So Velleman is (nearly) a Selfer.

I shall take my objections to the Selfer view to be objections to what Velleman (1996) says. So it is important that my objections to the Selfer view are not just objections to the view that you cannot survive any change in self, but are also objections to the thesis that you cannot survive those changes in self that involve your acquiring values, desires,

[2] Velleman thinks that memory gives first-personal access to a past point of view. But Velleman (1996, 74–6) denies that a person at a future time's remembering your current experiences implies that you now have first-personal access to that person's point of view at that future time. His explanation of this asymmetry involves double brain hemisphere transplant. And we have just seen another way to explain this asymmetry: you can remember what it was like to have very different values, desires, and projects at a past time, even if you were, at that past time, unable to know what it would be like to have your current values, desires, and projects.

and projects that you have never had before. And all the objections to the Selfer view to follow will be objections to that thesis.

Velleman says the following in a later work:

> . . . the word 'self' does not denote any one entity but rather expresses a reflexive guise under which parts or aspects of a person are presented to his own mind. This view stands in opposition to the view currently prevailing among philosophers—that the self is a proper part of a person's psychology, comprising those characteristics and attitudes without which the person would no longer be himself. I do not believe in the existence of the self so conceived. (2006, 1)

To have the 'same self' just is to be substantively alike with regard to values, desires, and projects (§I). So Velleman's being (nearly) a Selfer is consistent with his denying that the word 'self' denotes an entity. Nevertheless, the passage just quoted makes me think that Velleman would not be pleased with my conclusion that he is (nearly) a Selfer. For the passage just quoted makes me think that Velleman would not be pleased with my concluding that he is committed to the claim that there are some characteristics and attitudes—or at least some values, desires, and projects—that a person could not survive losing, at least not when they are replaced with values, desires, and projects that the person has never before had.

I admit that in this section I have not focused on the passage just quoted. And I have not even mentioned other aspects of Velleman's work that are relevant to these topics. Instead, I have focused on particular assertions and particular strands of argument in Velleman (1996) that—so I have argued—make him (nearly) a Selfer. So Velleman could rightly claim that in defending the conclusion that he is (nearly) a Selfer, I have not presented his overall views on personal identity. And Perry, Schechtman, and Whiting might each make a parallel claim about how I have defended the conclusion that each of them is a Selfer.

But no such claim is an objection to this chapter. For the goal of this chapter is not to explore in detail the overall views on personal identity offered by Velleman, Perry, Schechtman, or Whiting. As Chapter 3 (§V) emphasized, the goal is instead to consider answers to the Why Question that both compete with my answer and also are in terms of—or imply—psychological connectedness or psychological continuity.

And if the Selfer view is true, then every good answer to the Why Question both competes with my answer and also implies a sort of psychological connectedness (see §III). So the Selfer view is a threat to my answer.

Moreover, the Selfer view is not a merely hypothetical threat to my answer to the Why Question, present in logical space but not in the literature. We have seen that Velleman, Perry, Schechtman, and Whiting all endorse claims and defend strands of argument that imply the Selfer view. And Velleman says that the 'view currently prevailing among philosophers' is that there are certain 'characteristics and attitudes without which the person would no longer be himself'. If Velleman is right, then it seems that the Selfer view is not only present in the literature, but is prevalent.

III. The Same Self and Numerical Identity

Recall:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

The Selfer view has the result that every good answer to the Why Question implies having the same self. But the Selfer view is not itself an answer to the Why Question. So Selfers can and do answer the Why Question in a variety of ways. Perry answers in terms of a person's being likely to work on your projects at a future time (§I). Velleman answers in terms of having first-personal access to a person's point of view at a future time (§II). Schechtman answers in terms of having the same self-narrative (see Ch. 5, §I). And some Selfers will surely answer in terms of having the same self. But I do not think that any Selfer would answer in terms of numerical identity. Indeed, I think that every Selfer would deny that your being numerically identical with a person at a future time explains why you survive as that person. This is the first reason that the Selfer view competes with my answer to the Why Question, which answer is being numerically identical with.

I think that you could be numerically identical with a person who will exist at a future time even if the way you are now is not

psychologically connected to the way that person will be at that time. So you could be numerically identical with a person who will exist at a future time even if the way you are now is not substantively like the way that person will be at that time with regard to values, desires, and projects. So you could be numerically identical with a person at a future time even if that person will not have, at that time, the same self as you have.

Some Selfers agree. For example, Schechtman (1996, 26–50) defends the following argument. Numerical identity is transitive. Psychological connectedness is not transitive, not even the sort of psychological connectedness that is constituted by having the same self. So a person at the present time could—via the transitivity of identity—be numerically identical with a person at a future time even if that person will undergo a change in self between now and that future time.[3]

Whiting is not as explicit as Schechtman with regard to whether numerical identity could hold between a person at one time and a person at another time in the absence of the sort of psychological connectedness that is constituted by having the same self. But she does say:

> I doubt that I have any more reason to care about [Whiting existing at a future time] than about anyone else, if her experiences are related to mine only in the sense that they belong to the same immaterial soul . . . My general view is that the numerical identity of our present and future selves, for which sameness of soul has often been thought necessary, is irrelevant to justification of concern for our future selves. (1986, 547–8)

And Perry (1976) seems to think that you can persist as someone who will not have the same self as you now have. For he takes the following supposition quite seriously:

> Suppose I believe that tomorrow I will be struck by amnesia incurable in fact, though not in principle; that my character and personality will suddenly change. So I will hate what I now love, and work against what I now hope for. (1976, 84)

[3] Exact psychological connectedness is transitive. But no Selfer should take having the same self to be a matter of having exactly the same values, desires, and projects. For then the Selfer view would imply that no one can survive even a single small change with regard to values, desires, or projects. That would be a reductio of the Selfer view.

At any rate, I claim that it is possible for you to be numerically identical with a person at a future time even if that person will not have, at that time, the same self as you now have. My answer to the Why Question is being numerically identical with. So my answer implies that you can survive as a person at a future time even if that person will not have, at that time, the same self as you now have. That is, my answer to the Why Question implies that you can survive a change in self. The Selfer view just is the view that you cannot survive a change in self. This is the second reason that the Selfer view competes with my answer to the Why Question.

The Selfer view competes with my answer to the Why Question. That is, if the Selfer view is true, then my answer to the Why Question is not a good answer. Selfers might take this competition to be a reason to reject my answer, and so a reason to conclude that my argument for my answer in Chapter 2 (§I) must somehow go wrong. I myself take this competition to be a reason to deny the Selfer view. And I take my argument for my answer to the Why Question to be an argument for the falsity of the Selfer view.

I do not want to leave us at this impasse. So the rest of this chapter will argue that we should reject the Selfer view for reasons that do not depend on my answer to the Why Question, or on my defense of that answer. So no one should be a Selfer. So it would be a mistake to take the Selfer view's competing with my answer to be a reason to reject my answer or to be a reason to be suspicious of my defense of that answer.

IV. Growing Up

Astronauts and professional athletes aside, adults are not likely to work on the projects they started in kindergarten. Most adults do not act on the intentions they formed as first-graders. No preschooler's sense of self is built around being a philosopher, not even a preschooler who grows up to be an academic whose sense of self is built around being a philosopher. Prepubescents cannot know what it will be like to be postpubescent and, shall we say, restless along a particular dimension.

These few examples illustrate the obvious: the way that someone is as a child is not—with regard to values, desires, and projects—substantively psychologically connected to the way that that person will be as an adult. Indeed, a child's becoming an adult is as extreme a

change in values, desires, and projects as would be your becoming the sort of vampire discussed above. Selfers must deny that you would survive becoming a vampire. So Selfers must deny that a child will survive becoming an adult.

The change from child to adult is gradual. But Selfers must still deny that a child will survive becoming an adult. To see this, consider that even if you will change gradually into a vampire, Selfers must still deny that you will survive becoming a vampire. They must deny this for all the reasons presented in Sections I and II. Those reasons never turned on the claim that the change into a vampire—and the correlative change in self—was abrupt, as opposed to gradual.

I argued that Velleman is (nearly) a Selfer. That argument turned on the claim that one cannot know what it would be like to have a new and different self (§II). If that claim were false, then my argument for Velleman's being (nearly) a Selfer would fail. But even if that claim were false, Velleman's views would still imply that no child can survive becoming an adult. For Velleman's views would still imply that a child survives becoming an adult only if that child now has first-personal access to an adult's point of view at a future time. But no child now has first-personal access to an adult's point of view at a future time. This is because no child can now know what it would be like to be an adult. No child can know this even if an adult like you could, in at least some cases, know what it would be like to have a new and different self.

Selfers must deny that a child will survive becoming an adult. That is, Selfers must say that, for each child, there will never be an adult at some future time who will have, at that time, what matters in survival for that child. But Peter Pan is tragically stunted, as opposed to a tenacious survivor. Or so I say. For I claim that a child can survive becoming an adult, that is, an adult with the values, desires, and projects that are characteristic of an adult but not of a child. So I conclude that the Selfer view is false.

I do not think I need to defend the claim that a child can survive becoming an adult. For this claim should not be controversial. But I can defend it. Here is one defense. Persons can endure from childhood to adulthood; being numerically identical with is a good answer to the Why Question; therefore, a child can survive becoming an adult. But I shall set this defense aside. For I do not want this section's objection to the Selfer view to rely on my answer to the Why Question (see §III).

Here is another defense of the claim that a child can survive becoming an adult. If a person's experiences at a future time will be good (or bad) for the you of right now, then that person will have, at that time, what matters in survival for you (Ch. 1, §§II–III). Premise: it will be good for some toddler to enjoy a full life as the happy adult she will become. So that toddler will survive as that adult. So a toddler can survive becoming an adult. So a child can survive becoming an adult.[4]

The premise that it will be good for some toddler to enjoy a full life as the happy adult she will become should not be controversial. But here is a defense of that premise. Virtually everyone who is or has been the parent of a toddler has hoped, for that toddler's own sake, that that toddler will enjoy a full life as a happy adult. I deny that every time a parent has hoped for this, that parent has hoped for the impossible. So it will be good for some toddler to enjoy a full life as the happy adult she will become. So a toddler can survive becoming an adult.[5]

The Selfer view implies that no child can survive growing up. And that is not the only false implication that the Selfer view has with regard to normal human lives. Teenagers differ significantly in values, desires, and projects from the forty-year-olds that most of them will become. And many old women and old men have undergone at least one change in self since their thirtieth birthday. And so on. I say that

[4] Jeff McMahan agrees that a future experience will be good for you just in case whoever will have that future experience will then have what matters in survival for you. But we draw different conclusions from this. For McMahan (2002, 78) says that since no adult will have what matters in survival for a late-term fetus, the death of a fetus (and so the loss of potential goods that would be enjoyed in a long life) is 'not a terrible tragedy, at least not for the fetus itself'. McMahan is explicit that the same goes for infants. I suppose he must say the same about small children, since his reason for saying what he does about fetuses and infants is that 'psychological unity is among the prudential unity relations' (McMahan, 2002, 78).

[5] As Chapter 1 (§II) emphasized, its being appropriate for you to first-personally anticipate a future experience does not imply that you do in fact first-personally anticipate that experience. The same goes for its being appropriate for you to have future-directed self-interested concern with regard to an experience. And, of course, children do not first-personally anticipate, or have future-directed self-interested concern with regard to, many of the experiences that each will have as an adult. But the experiences that a child will have as an adult will be good (or bad) for that child. So it is appropriate for that child to have future-directed self-interested concern with regard to, and so first-personally anticipate, those experiences.

we often survive evolving over the decades in these ways. The Selfer view implies that no one ever survives any of this. So I conclude that the Selfer view is false.

This chapter is focused on Selfers. So it is focused on those who take survival to require the sort of substantive psychological connectedness that is constituted by having (enough of) the same values, desires, and projects. But anyone who takes survival to require substantive psychological connectedness of any sort is subject to the worries raised in this section.

For example, recall Leibniz's view that a future person will have what matters in survival for you only if that person will have memories of your current life (Ch. 1, §§I–II). Adults do not remember being toddlers, at least not typically. So Leibniz's view implies that, at least typically, toddlers will not survive becoming adults.

Here is another example, from David Lewis:

> Consider Methuselah. At the age of 100 he still remembers his childhood. But new memories crowd out the old. At the age of 150 he has hardly any memories that go back before his twentieth year . . . As he grows older he grows wiser; his callow opinions and character at 90 have vanished almost without a trace by age 220, but his opinions and character at age 220 also have vanished almost without a trace by age 350. He soon learns that it is futile to set goals for himself too far ahead. At age 120 he is still somewhat interested in fulfilling the ambitions he held at age 40; but at age 170 he cares nothing for those ambitions, and it is beginning to take an effort of will to summon up an interest in fulfilling his aspirations at age 80. And so it goes. (1976, 30)

Lewis says that 'what matters in survival is literally *identity*' (1976, 20).[6] Lewis (1976) takes substantive psychological connectedness to be necessary for identity. So he takes it to be necessary for survival. So he concludes that Methuselah at age 80 does not survive as, for example, Methuselah at age 350.

Lewis wants to draw a distinction here between 'us short-lived' creatures and Methuselah. Lewis says:

[6] Lewis (1976, 17) also says: '*what matters in survival is mental continuity and connectedness*'. One of the main points of Lewis (1976) is to argue that these two claims about what matters in survival are consistent. That argument explicitly trades on the claim that persons perdure, and so do not endure.

> We sometimes say: in later life I will be a different person. For us short-lived creatures, such remarks are an extravagance . . . For Methuselah, however, the fading-out of personal identity looms large as a fact of life. It is incumbent on us to make it literally true that he will be a different person after one and one-half centuries or so. (1976, 30)

But I deny that this distinction can be drawn. The reasons Lewis gives for concluding that Methuselah will not survive for more than one-and-a-half centuries are equally good—that is, equally bad—reasons for concluding that no child will survive becoming an adult.

V. Other Transformations

Case One: You learn that someone who is not a Nazi will, over the next twenty years, gradually become a Nazi. You are disturbed and saddened. You learn next that that someone is you. Then, as Perry (§I) would put it, a whole new set of emotions rises within your breast: you dread becoming a Nazi; you are horrified not just by Nazi values, desires, and projects, but also by your coming to have those values, desires, and projects; and so on. You thereby first-personally anticipate, and have future-directed self-interested concern with regard to, experiences that you will have as a Nazi.

I do not think that you can now know what it would be like to be a Nazi (cf. §II). But in Case One you do first-personally anticipate the experiences you will have as a Nazi. So I reject Velleman's characterization of first-personal anticipation. For, as we saw in Section II, Velleman's characterization implies that you can first-personally anticipate an experience a person will have at a future time only if you can know what it would be like to be that person, having that experience, at that time.

Or consider Mary, who has been in a black-and-white room her whole life (cf. Jackson, 1982). She knows that she will soon leave this room and have the experience of seeing red for the first time. She is now looking forward to seeing red, is now feeling excitement about seeing red, is now wondering what it will be like for her to see red, and so on. So Mary is now both first-personally anticipating, and also having future-directed self-interested concern with regard to, seeing red. This is all consistent with the claim that Mary cannot now know what it would be like to see red.

Here is Case Two: You are not a Nazi. You learn that you will be annihilated twenty years from now. A Nazi will then be created at the same instant in your place. This is bad news for you. But the bad news for you is that you will be annihilated. The bad news for you is not that you will become a Nazi. So you do not dread becoming a Nazi, nor are you horrified by your coming to have Nazi values, and so on. So you do not first-personally anticipate, or have future-directed self-interested concern with regard to, experiences that will be had by the Nazi who will replace you.

You are not now a Nazi. So the values, desires, and projects that 'make you the person you are' are not those of a Nazi. So no future Nazi will be—with regard to values, desires, and projects—substantively psychologically connected to the way you are now. So the Selfer view implies that no future Nazi could have what matters in survival for you. So the Selfer view implies that it is no more appropriate for you to first-personally anticipate, with horror, having the experiences of a Nazi in Case One than it is to first-personally anticipate this in Case Two.

If I were to discover that I was going to become a Nazi, this would horrify me in ways that the discovery that I was going to be annihilated would not, not even if I also discovered that a Nazi was going to be created upon my annihilation. And I do not think my being horrified in this way would be based on a mistake. So I think that in Case One, it is appropriate for you to first-personally anticipate, and have future-directed self-interested concern with regard to, the experiences that you will have as a Nazi. So I think that the Selfer view is false.

Look at it this way. The experiences of a future (conscious) person will be bad (or good) for you—for the you of right now—just in case you will survive as that future person (Ch. 1, §§II–III). So it will be bad for you to have a future Nazi's values, desires, and projects just in case you will survive as that Nazi. So Selfers must deny that coming to have Nazi values, desires, and projects would be bad for you. Indeed, Selfers must deny that becoming a Nazi would be bad for you. Again, I think that the Selfer view is false.

The Selfer view is not consistent with the claim that becoming a Nazi would be bad for you. More carefully, the Selfer view is not consistent with the claim that becoming a Nazi would be distinctively bad for you. But the Selfer view is consistent with—and even

implies—that becoming a Nazi would be bad for you in the way that being annihilated would be bad for you. For just as Leibniz says that forgetting all you have been is the practical equivalent of annihilation (Ch.1, §I), so Selfers should say that undergoing a change in self is the practical equivalent of annihilation. But I do not think that this helps the Selfer view. Instead, as we shall see, this is yet another reason to reject the Selfer view.

Thomas Nagel says:

> If we are to make sense of the view that to die is bad, it must be on the ground that life is a good and death is the corresponding deprivation or loss, bad not because of any positive features but because of the desirability of what it removes.
>
> (1970, 75) (see also Nagel, 1986, 224; Feldman, 1991; Bradley, 2009, 47–60; and Broome, 2013)

Suppose Nagel is right. Then annihilation—like death—would be bad for you only in that it deprives you of the goods you would otherwise enjoy. I deny that becoming a Nazi would be bad for you only in that it would deprive you of enjoying certain goods. So I deny that becoming a Nazi would be bad for you in the way that being annihilated would be bad for you.

Another view is that death will not be bad for you. This view is motivated by the following well-known line of reasoning: nothing will be bad for you unless you will be there to experience it; and you will not be there to experience anything after you have died.[7] Suppose this reasoning shows that death will not be bad for you. Then being annihilated would not be bad for you either. Then it is false that becoming a Nazi would be bad for you in the way that being annihilated would be bad for you, that is, not bad for you at all. For, again, becoming a Nazi would be bad for you.

We have considered only two views about the badness (or not) of annihilation. But I do not think that there is any account of how being annihilated would (or would not) be bad for you that can plausibly be taken to also be an account of how becoming a Nazi would be bad for you. The Selfer view implies that becoming a Nazi would be bad

[7] See Epicurus, 'Letter to Menoeceus' 124–5 [1926, 85] and Lucretius, *On the Nature of the Universe* 3.843–3.862 [1994, 88].

for you only in the way that being annihilated would be bad (or not) for you. So, again, I think that the Selfer view is false.

I have focused on a tragic transformation. But transformations can be beautiful. You are now a Nazi. (Sorry.) But in twenty years you will not only be a non-Nazi, you will be loving and wise. I think that it will be good for you—for the you of right now—to become loving and wise. But Selfers must deny this. Selfers must say that this will be bad for you to the extent—and in the way—that being annihilated would be bad for you.

So the Selfer view implies that it would be bad for a Nazi to become a non-Nazi to the extent—and in the way—that being annihilated would be bad for that Nazi. And the Selfer view implies that it would be bad for a non-Nazi to become a Nazi to the extent—and in the way—that being annihilated would be bad for that non-Nazi. So a Nazi's becoming a non-Nazi who is loving and wise is just as bad for that Nazi as a non-Nazi's becoming a Nazi is bad for that non-Nazi. Or so the Selfer view implies. Again, I think that the Selfer view is false.

In George Orwell's *1984*, Winston Smith has values, desires, and projects that are not approved of by the Party. So the Party's Thought Police forcibly 'reintegrate' Smith. They make Smith have (via torture) the values, desires, and projects that are approved of by the Party. Reintegration causes a change in self (cf. Barnes, 2015).

The Thought Police harmed Smith by torturing him. But they also harmed Smith by replacing the values, desires, and projects that Smith had with worse values, desires, and projects. They harmed him by making him into a worse sort of person. If you agree that Smith was harmed in this way, then you should conclude that it was bad for the 'pre-reintegrated' Smith to be reintegrated. But then Smith survived reintegration. So the Selfer view is false.

Or suppose you somehow stunt a child's psychological growth so that, over the coming decades, she will not change in values, desires, or projects. Then you have grievously harmed that child. But the Selfer view implies that you have, instead, protected that child from the practical equivalent of annihilation. Again, the Selfer view is false.

The Selfer view manages to be both implausibly optimistic and implausibly hopeless at the same time. Here are two examples of implausible Selfer optimism: it is never appropriate for you to

first-personally dread tragic changes in self, and no change in self can be bad for a person, other than being bad for a person in the way that being annihilated would be bad (or not) for a person. Here are two examples of implausible Selfer hopelessness: it is never appropriate for you to first-personally anticipate positive changes in self, and no change in self can be good for a person. I conclude, again, that the Selfer view is false.

This section's objections to the Selfer view can be turned into objections to other instances of the claim that survival requires psychological connectedness. For example, suppose that it would be bad for you—for the you of right now—to lose all your memories tomorrow. That is, suppose that this would be bad for you in a way other than how being annihilated would be bad (or not) for you. Then you would survive losing your memories. Then we have a new objection to Leibniz's view that a future person will have what matters in survival for you only if that person will have memories of your current life (Ch. 1, §§I–II).

VI. Conclusion

The Selfer view says that a person at a future time will have, at that time, what matters in survival for you only if the way that person will be at that time is—with regard to values, desires, and projects—substantively like the way you are now. That is, the Selfer view says that you cannot survive a change in self. We have seen that the Selfer view is false. So any good answer to the following must allow for you to survive a change in self:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

My answer is being numerically identical with. You can be numerically identical with a person at a future time even if the way that person will be at that time is—with regard to values, desires, and projects—not substantively like the way you are now. So my answer allows for you to survive a change in self.

5

The Same Self-Narrative

Consider:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

I shall briefly discuss an answer to this question that is in terms of narrative, but that does not imply psychological connectedness or psychological continuity (§IV). But I shall not focus on this sort of answer. This is because I want to focus on the most serious competitors to my answer to the Why Question. And I take the most serious competitors to be in terms of—or at least to imply—psychological connectedness or psychological continuity (Ch. 3, §V).

The narrative-based answers to the Why Question that I shall focus on are in terms of psychological connectedness, and in terms of psychological continuity (which is a chain of overlapping instances of psychological connectedness). Chapter 6 considers an answer in terms of narrative-based psychological continuity. But this chapter focuses on an answer in terms of narrative-based psychological connectedness. In particular, this chapter focuses on an answer in terms of being alike with regard to 'self-narrative'. As we shall see, this chapter thereby continues the discussion of the Selfer view that began in Chapter 4.

I. The Self-Narrative Account

Marya Schechtman says:

The cornerstone of the narrative self-constitution view is the claim that a person's identity is created by a self-conception that is narrative in form. Most

Self and Identity. Trenton Merricks, Oxford University Press. © Trenton Merricks 2022.
DOI: 10.1093/oso/9780192843432.003.0006

broadly put, this means that constituting an identity requires that an individual conceive of his life as having the form and the logic of a story—more specifically, the story of a person's life—where 'story' is understood as a conventional, linear narrative. (1996, 96)

Let a person's *self-narrative* be the story of that person's life.

Schechtman says: 'the narrative self-constitution view is the claim that a person's identity is created by a self-conception that is narrative in form'. She does not thereby say that a self-narrative creates a person's numerical identity, whatever that would mean. Rather, according to Schechtman, a self-narrative creates a different kind of identity, namely:

. . . the kind of identity that is at issue in an 'identity crisis' and not the logical relation of identity . . . The notion to which I appeal here is the general use of the term 'identity' to refer to the set of characteristics each person has that makes her the person she is. In an identity crisis, a person is unsure about what those defining features are, and so is unsure of his identity. (1996, 74)

So Schechtman says that the values, desires, and so on that 'make you the person you are' are those that are included in your self-narrative. Again, Schechtman (1996, 94) says: ' . . . a person's *identity* . . . is constituted by the content of her self-narrative, and the traits, actions, and experiences included in it are, by virtue of that inclusion, hers'. And Schechtman thinks that your self-narrative—and so your 'identity'—is to some extent up to you.

Schechtman (1996, 99) says: 'the goal of the narrative self-constitution view is to capture the intuitive relation between personal identity and the four features'. One of those 'four features' is 'self-interested concern' (Schechtman, 1996, 2). And Schechtman takes having the same identity of the sort just noted to be the sort of 'personal identity' that makes self-interested concern appropriate.

So Schechtman (1996) would say that it is appropriate for you to have self-interested concern with regard to a person's experiences at a future time just in case your identity now is the same as that person's identity at that future time. She thinks you and a person at a future time have the same identity in virtue of having the same self-narrative. So Schechtman is committed to the claim that it is appropriate for you to have self-interested concern with regard to a person's experiences at a future time just in case—and because—the self-narrative that you

now have will be had by that person at that future time. And this claim implies a parallel claim about appropriate first-personal anticipation (Ch. 1, §II).

Recall:

> *The What Question*: What is it for a person at a future time to have (at that time) what matters in survival for you?

The answer: its being appropriate for you to first-personally anticipate the experiences that that person will have at that future time; and if that person will have good (or bad) experiences at that time, its being appropriate for you to have future-directed self-interested concern with regard to those experiences (Ch. 1, §II).

So Schechtman is committed to the following two claims. First, a person at a future time will have, at that time, what matters in survival for you if and only if that person will then have the self-narrative that you now have. Second, you survive as that person at that future time because that person will then have the self-narrative that you now have; in other words, having the same self-narrative is what explains your surviving as that person. Let the *Self-Narrative Account of survival* be these two claims. That account answers:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

That account's answer is having the same self-narrative.

In what follows, I shall proceed as if Schechtman takes *having the same self-narrative* to be a way of being substantively psychologically alike. I have two reasons for proceeding in this way. The first reason is that Schechtman is a Selfer. Schechtman thinks that a person at a future time will have, at that time, what matters in survival for you only if that person will then have the values, desires, and projects that now 'make you the person you are' (see Ch. 4, §I). And Schechtman thinks that it is the values, desires, and projects that are included in your self-narrative that make you the person you are.

Schechtman's being a Selfer makes perfect sense if having the same self-narrative is a way of being substantively psychologically alike. For then Schechtman can say that a person at a future time who will have the same self-narrative as you have will thereby be like you with regard

to the values, desires, and projects that make you the person you are. Moreover, Schechtman can say that you and that person will be alike in that way because you are alike with regard to having a self-narrative that 'includes' those values, desires, and projects.

On the other hand, suppose—just for the sake of argument—that having the same self-narrative is not a matter of being substantively psychologically alike. Then having the same self-narrative does not imply being substantively psychologically alike, not even with regard to values, desires, and projects. So then the Self-Narrative Account has the result that you could survive as a person at a future time—in virtue of having the same self-narrative as that person—even if that person will not have the values, desires, and projects that you now have. But then the Self-Narrative Account is not consistent with the Selfer view. And then the Self-Narrative Account is not consistent with the following, which is central to Schechtman's (1996) narrative self-constitution view: it is appropriate for you to have future-directed self-interested concern with regard to a person's experiences at a future time just in case that person will have your 'identity', that is, will have the values, desires, and projects that 'make you the person you are'.

I have a second reason for proceeding as if Schechtman takes *having the same self-narrative* to be a way of being substantively psychologically alike. This reason begins with how Schechtman describes having a self-narrative. She says:

> The way in which we have autobiographical narratives on [the narrative self-constitution view] is cashed out mostly in terms of the way in which an implicit understanding of the ongoing course of our lives influences our experience and deliberation. (2014, 101)

Schechtman (1996, 113) also emphasizes that a self-narrative is the 'lens through which we filter our experience and plan for actions'. And, as we saw above, Schechtman (1996, 96) thinks that to have a self-narrative is to conceive of your life as having the form and the logic of a story.

But then it seems that your having the same self-narrative as a person at a future time just is your implicitly understanding the ongoing course of your life in the way that that person will implicitly understand, at that future time, the ongoing course of her or his life. Or it is the 'lens' through which you view your experience and plan for

action being the same as that person's 'lens' at that future time. Or it is the story you conceive of your life as having being the same as the story that that person will, at that future time, conceive of his or her life as having. All this implies that you have the same self-narrative as a person at a future time just in case the way you are now is psychologically like the way that person will be in a particular way.

For the above two reasons I shall proceed as if Schechtman says that having the same self-narrative is a way of being substantively psychologically alike. But I admit that some of what Schechtman says does not fit with this interpretation. For example, she says:

> I call a person's underlying psychological organization a *self*-narrative because it is not simply a static set of facts about him, but rather a dynamic set of organizing principles, a basic orientation through which, with or without conscious awareness, an individual understands himself and his world. These implicitly organizing principles are not simply a collection of features, but a continually developing interpretation of the course of one's own trajectory through the world. (1996, 115–16)

Schechtman here seems to take a self-narrative to be dynamic. That is, she seems to take a self-narrative to 'continually develop' in a way that is consistent with your having the same self-narrative as a person at a future time even if you fail to be substantively psychologically like that person at that time.

But I do not think we should read Schechtman as taking a self-narrative to be dynamic. This is mainly because of the above two reasons for taking Schechtman to say that having the same self-narrative is a way of being substantively psychologically alike. But there is one other reason. This other reason starts by pointing out that if a self-narrative is dynamic, then having the same self-narrative seems to be a species of psychological continuity, as opposed to psychological connectedness. That is, your having the same self-narrative as a person at a future time seems to be a matter of your being related by overlapping instances of psychological connectedness to that person at that time. (Each such instance of connectedness is constituted by having a similar 'interpretation of the course of one's own trajectory through the world'.)

Psychological continuity is transitive. So if a self-narrative is dynamic, then having the same self-narrative is transitive. So if Schechtman takes a self-narrative to be dynamic, then Schechtman

answers the Why Question in terms of a transitive relation: having the same (dynamic) self-narrative. But Schechtman argues at length that the answer to (what I call) the Why Question cannot be transitive. Indeed, one of Schechtman's main reasons for denying that numerical identity is a good answer to the Why Question is that numerical identity is transitive. Schechtman devotes an entire chapter to this reason, a chapter called 'The Problems of Logical Form' (1996, 26–50) (see also Schechtman (1996, 78–89) and Ch. 4, §III).

At any rate—and regardless of how you interpret Schechtman (1996)—let us all understand *having the same self-narrative* to be a matter of being substantively psychologically alike in the way described above. Then having the same self-narrative is a species of psychological connectedness. Then the Self-Narrative Account of survival answers the Why Question in terms of a species of psychological connectedness. So this chapter explores an answer to the Why Question in terms of psychological connectedness, an answer that competes with my answer (see §III).

II. The Same Self-Narrative and the Same Self

The Self-Narrative Account of survival implies that a person at a future time will have, at that time, what matters in survival for you only if that person will then have the self-narrative that you now have. So that account implies that you cannot survive losing your self-narrative. Presumably, if you lose (enough of) the values, desires, and projects included in your self-narrative, you thereby lose your self-narrative. So the Self-Narrative Account implies that you cannot survive losing the values, desires, and projects included in your self-narrative. Suppose that the values, desires, and projects included in your self-narrative are those that 'make you the person you are'. Then the Self-Narrative Account implies that you cannot survive losing the values, desires, and projects that 'make you the person you are'. Then the Self-Narrative Account implies the Selfer view.

Recall that the vampires say: 'Life has meaning and a sense of purpose now that it never had when I was human. I understand Reality in a way I just couldn't before' (Ch. 4, §I). To acquire the values, desires, and projects of a vampire is thereby to acquire a vampire's understanding of the course of one's life and a vampiric

lens for filtering experiences and planning for action. So your acquiring the values, desires, and projects of a vampire would result in your acquiring a new self-narrative. More generally, it seems that a change in (enough of) your values, desires, and projects results in a change in self-narrative. That is, a change in self results in a change in self-narrative. So if you cannot survive a change in self-narrative, then you cannot survive a change in self. The Self-Narrative Account of survival implies that you cannot survive a change in self-narrative. So that account implies that you cannot survive a change in self. Again, the Self-Narrative Account implies the Selfer view.

Some deny that persons in general have self-narratives (see, e.g. Strawson, 2004). They will deny that a change in self results in a change in self-narrative. This does not matter. For their denials do not undermine the following line of reasoning. If the Self-Narrative Account is true, then having the same self-narrative as a person at a future time is necessary for surviving as that person. So that account implies that you can survive only if you have a self-narrative. Persons in general can survive. So if the Self-Narrative Account is true, then persons in general have self-narratives. In other words, the Self-Narrative Account implies that persons in general have self-narratives. So the Self-Narrative Account implies that a change in self results in a change in self-narrative. So if the Self-Narrative Account is true, you cannot survive a change in self. Again, the Self-Narrative Account implies the Selfer view.

The Self-Narrative Account says that your having the same self-narrative as a person will have at a future time is sufficient for your surviving as that person at that time. Defenders of that account will want their view to be consistent with the fact that, typically, at most one person at a future time will have, at that time, what matters in survival for you. (Teletransportation gone awry is atypical.) So they should say that, typically, at most one person at a future time will have, at that time, the same self-narrative as you now have. So they should say that the sort of psychological connectedness that constitutes having the same self-narrative is substantive. So I take the Self-Narrative Account to imply not just that persons in general have self-narratives, but that they have quite detailed self-narratives.

The Self-Narrative Account implies the Selfer view. So any argument for the falsity of the Selfer view is thereby an argument for the

falsity of the Self-Narrative Account. Chapter 4 (§§III–V) presented several arguments for the falsity of the Selfer view. So I conclude that the Self-Narrative Account is false. And I shall raise more objections to that account below, objections that target that account directly, as opposed to by way of targeting the Selfer view. But some of those objections will be similar to Chapter 4's objections to the Selfer view. This is because the Self-Narrative Account—like the Selfer view—takes substantive psychological connectedness to be necessary for survival.

III. The Same Self-Narrative and Numerical Identity

I say that being numerically identical with is a good answer to:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

And I think that it is possible for you to be numerically identical with a person who will exist at a future time even if the way you are now is not substantively psychologically connected to the way that person will be at that time. So I conclude that it is possible for a person at a future time to have, at that time, what matters in survival for you even if that person will fail to have, at that time, the same self-narrative as you have now. The Self-Narrative Account of survival implies that this is not possible. So the Self-Narrative Account competes with my answer to the Why Question.

There is a second reason that I take the Self-Narrative Account of survival to compete with my answer to the Why Question. I assume that the Self-Narrative Account implies that the only good answers to the Why Question are in terms of having the same self-narrative or (perhaps) in terms of related sorts of psychological connectedness, such as having the values, desires, and projects that are included in your self-narrative. So the Self-Narrative Account implies that it is false that being numerically identical with is a good answer to the Why Question. So that account competes with my answer.

The Self-Narrative Account of survival competes with my answer to the Why Question. So if my answer to the Why Question is a good answer, then the Self-Narrative Account is false. So my defense of my

answer to the Why Question is thereby a defense of the falsity of the Self-Narrative Account (Ch. 2, §I). So I conclude that the Self-Narrative Account of survival is false.

Again, the Self-Narrative Account of survival competes with my answer to the Why Question. So if the Self-Narrative Account is true, then my answer to the Why Question is not a good answer. So you might take the Self-Narrative Account's competing with my answer to the Why Question to be a reason to reject my answer, and so a reason to conclude that something or other must have gone wrong in my defense of that answer. But that would be a mistake.

Any reason to reject the Selfer view that does not presuppose my answer to the Why Question is thereby a reason to reject the Self-Narrative Account that does not presuppose my answer to the Why Question (see §II). And Chapter 4 (§§IV–V) presents a number of such reasons. Moreover, Sections IV and V below present more reasons to reject the Self-Narrative Account, reasons that do not presuppose my answer to the Why Question. So Chapter 4 (§§IV–V) and Sections IV and V show that it would be a mistake to take the Self-Narrative Account's competing with my answer to the Why Question to be a reason to reject my answer, or a reason to conclude that something or other must have gone wrong in my defense of that answer.

IV. Growing Up Redux

Consider this disjunctive claim: a small child has an implicit understanding of the ongoing course of that child's life; or a small child has a lens through which that child filters experiences and plans for action; or a small child conceives of his or her own life as having the form and logic of a story; or a small child understands herself or himself as a character in a narrative who has certain values, desires, and projects. Add disjuncts for other accounts of what it might be to have a self-narrative.

Suppose that this disjunctive claim is false. Then small children do not have a self-narrative. So no small child has the same self-narrative as will an adult at a later time. Suppose that this disjunctive claim is true. Suppose, for example, that small children have an implicit understanding of the ongoing course of their respective lives. Adults

do not implicitly understand the ongoing course of their respective lives in the way that small children do. So no small child has the same implicit understanding of the ongoing course of his or her life as will an adult at a later time. More generally—and if this disjunctive claim is true—no small child has the same self-narrative as will an adult at a later time.

The above disjunctive claim is either true or false. Either way, we should conclude that no small child has the same self-narrative as will an adult at a later time. The Self-Narrative Account of survival says that a person at a future time has, at that time, what matters in survival for you if and only if the self-narrative that you now have is the same as the self-narrative that that person will have at that future time. So the Self-Narrative Account of survival implies that no small child will survive growing up.

Premise: small children often survive becoming adults. Also: it will be good for some small child to enjoy a full life as the happy adult she will become; so some small children will survive growing up (see Ch. 4, §IV). Again, the Self-Narrative Account of survival implies that no small child will survive growing up. So I conclude that the Self-Narrative Account is false.

Schechtman (1996, 146) says: 'Infants do not narrate their lives.' And Schechtman adds:

> Notably, infants and very young children also do not have in their repertoire the kinds of psychological activities that require persistence through time and are definitive of persons. (1996, 146)

Schechtman denies that very young children are persons. She also thinks that they do not have self-narratives and so do not have 'identities'. So I think Schechtman would say that no very young child can survive growing up.

Suppose you agree with Schechtman. You should still reject the Self-Narrative Account of survival. For that account does not go wrong only when it comes to very young children. Each teenager's lens through which he or she filters experiences and plans for action is—one hopes—not the same as the lens of the forty-year-old that that teenager will become. Most old women and old men understand the ongoing course of their lives in a different way than they did when each was thirty years old. And so on. I do not think that evolving over the

decades with regard to how you understand your life is the practical equivalent of annihilation. On the contrary, it can be just more growing up. Again, I conclude that the Self-Narrative Account is false. This concludes this section's discussion of the Self-Narrative Account of survival.

The rest of this section considers a different account of survival. This account invokes self-narrative but is tailor-made to allow a small child to survive growing up. Considering this account will allow me to defend the claim that not only is the Self-Narrative Account exactly as understood above inconsistent with a child's surviving growing up, but so is the basic idea that self-narrative is what delivers survival.

David DeGrazia says:

> It doesn't matter that one can't remember being born and might have trouble anticipating a state of severe dementia. One knows on the basis of others' testimony and everyday biological and medical knowledge that one was born and might someday become demented. Thus the past event is appropriated into one's inner story, and the possible future state is appropriated as a possible continuation of the story. (2005, 83)

Suppose that your self-narrative (your 'inner story') includes your having been a particular person at an earlier time. Add that this implies that that person at that earlier time has survived as you. Add further that if your self-narrative includes your being a particular person at a later time, then you will survive as that person at that later time. All this constitutes a self-narrative-invoking account of survival. Let this be the *Kid-Friendly Account of survival*.

Pretend that it is part of my self-narrative that I was once Emperor of France, was defeated at Waterloo, and so on. Add the Kid-Friendly Account of survival. Then it seems that we get the result that Napoleon has survived as me. This is a bad result. So let us assume that we can avoid this bad result by taking the Kid-Friendly Account to include a 'reality constraint'.

DeGrazia endorses a reality constraint. In particular, DeGrazia thinks that your including a past (or future) person in your self-narrative delivers survival only if you are numerically identical with that past (or future) person (DeGrazia, 2005, 114).

Schechtman also endorses a reality constraint. She says:

> . . . the facts with which an identity-constituting narrative must cohere obviously cannot be facts about persons *per se*, or the narrative self-constitution

> view would be viciously circular. On this view the kinds of facts to which a narrative must be responsible are thus not facts about persons, but facts about human beings and their environments. (1996, 120)

Schechtman (1996, 120–1) thinks that your including a past (or future) person in your self-narrative delivers survival only if you are the same human being as that person.

An adult can have an implicit understanding of the ongoing course of his or her life that includes having been this or that particular small child. This implicit understanding can accommodate the reality constraints just considered. For that adult can be numerically identical with, and the same human being as, that child. So the Kid-Friendly Account is aptly named. That is, the Kid-Friendly Account allows for a small child to survive becoming an adult. But I do not think that anyone who endorses the basic idea that self-narrative is what delivers survival should endorse the Kid-Friendly Account. A couple of examples should make this clear.

First example: You will be bitten and then transform into a vampire. Let this transformation be a paradigm case of acquiring a new and different self-narrative. So your transformation will be partly constituted by your immediately acquiring a new lens through which you view your experiences and plan for action, and so on. After transformation, the vampire will say: 'I am so glad I was bitten. Before that happened, I was a miserable piece of work: implicitly understanding the ongoing course of my life in terms of mundane projects, bourgeois values, and prosaic desires.'

Becoming this sort of vampire is a paradigm case of acquiring a new and different self-narrative. So I think that anyone attracted to the basic idea that self-narrative is what delivers survival should deny that you will survive as a vampire in this example. So they should deny that a future vampire's including the you of today as part of that vampire's implicit understanding of the ongoing course of that vampire's life (etc.) implies that you will survive as that vampire.

Look at it this way. That vampire will remember what she or he once was. So that vampire will remember being the you of today. So that vampire will include being the you of today as part of that vampire's story. (That story accommodates any reasonable reality constraint, since it is you who will persist as that vampire.) But an answer to the Why Question in terms of self-narrative is supposed to

require more for you to survive as a future vampire than that vampire's remembering being the you of today. More generally, to say that memory delivers survival is not thereby to say that self-narrative is what delivers survival.

Second example: A new genetic test has revealed that I shall suffer dementia in twenty years, dementia so severe that I shall no longer have a self-narrative. Part of my current implicit understanding of the ongoing course of my life now includes my being someone who will thus suffer. (That implicit understanding accommodates any reasonable reality constraint, since it is I who shall suffer.) Schechtman (1996, 146–50) would deny that the demented me at some future time will have, at that time, what now matters in survival for me. For example, Schechtman says:

> . . . the most common actual cases of human beings without narratives are the ones already discussed in this chapter—infants and sufferers of dementia . . . . to contemplate a descent into a second infancy such as that present in late-stage Alzheimer's is, as has been observed repeatedly, to contemplate the loss of oneself—a personal death. (1996, 150)

Others who think that self-narrative is what delivers survival should join Schechtman. For the claim that self-narrative is what delivers survival should imply that no one will survive as a person at a future time who, at that future time, will have no self-narrative.

The Kid-Friendly Account implies that you will survive as a vampire in the first example. And the Kid-Friendly Account implies that I shall survive as someone with no self-narrative in the second example. So those attracted to the basic idea that self-narrative is what delivers survival should reject the Kid-Friendly Account. This is reason enough to set that account aside for the rest of this chapter, since this chapter is interested in the idea that self-narrative delivers survival. But there is one more reason to set the Kid-Friendly Account aside, a reason to set it aside for the rest of this book.

Suppose that I learn that the way I shall be in twenty years will be neither psychologically connected to nor psychologically continuous with the way I am now. Add that my 'inner story' now includes my being a person who, in twenty years, will not be psychologically connected to, or continuous with, the way I am now. (And add that that person will be the numerically same human being as am I.) The

Kid-Friendly Account implies that I shall survive as that person twenty years from now, a person who will then be neither psychologically connected to, nor continuous with, the way I am now. So the Kid-Friendly Account does not lead to an answer to the Why Question that is in terms of—or even implies—any sort of psychological connectedness or psychological continuity.

Or instead suppose—just for the sake of argument—that the way you are now is neither psychologically connected to nor psychologically continuous with the way you were when you were two years old. Add that your current self-narrative includes your being that two-year-old child. The Kid-Friendly Account implies that that child survived as you. So the Kid-Friendly Account does not lead to an answer to the Why Question that is in terms of—or even implies—any sort of psychological connectedness or psychological continuity.

I answer the Why Question in terms of numerical identity (Ch. 2, §I). My answer is not in terms of—and might not imply—any sort of psychological connectedness or psychological continuity. I take the most serious competitors to my answer to the Why Question to be in terms of—or to imply—psychological connectedness or psychological continuity (Ch. 3, §V). These are the competitors I want to explore. The Kid-Friendly Account is not, and does not lead to, one of these competitors. So I am going to set the Kid-Friendly Account aside.

V. Other Changes in Self-Narrative

Suppose that I shall suffer self-narrative-erasing dementia in twenty years. Add that I know this. So I first-personally anticipate (dread) the experiences that I shall then have. And I have future-directed self-interested concern with regard to those experiences. All of this seems to be appropriate. And if this is appropriate, then I shall survive as the person who will have those experiences, but no self-narrative, in twenty years. So I shall survive as a person with no self-narrative.

I believe that acquiring self-narrative-erasing dementia will be bad for me, that is, bad for the me of right now. Moreover, I do not think that this will be bad for me in the way in which being annihilated would be bad (or not) for me. So I shall survive as the person who will be thus demented (Ch. 1, §§II–III). So, again, I shall survive as a person with no self-narrative.

So the Self-Narrative Account of survival is false. For that account implies that no one will survive as someone with no self-narrative. This is because that account says that a person at a future time will have, at that time, what matters in survival for you if and only if that person will then have the self-narrative that you now have.

You are not now a Nazi. Your current self-narrative is not heinous. It would be bad for you to replace your current self-narrative with a heinous self-narrative that takes your life to be a narrative quest to bring back the Third Reich, to come to implicitly understand the ongoing course of your life in light of Nazi values and projects, and so on. And all this would be bad for you in a way other than how being annihilated would be bad (or not) for you. But the Self-Narrative Account falsely implies that this would be bad (or not) for you only in the way in which being annihilated would be bad (or not) for you (see Ch. 4, §V). For that account implies that you cannot survive exchanging one self-narrative for another.

Suppose you currently have the above heinous self-narrative. It would be good for you to replace that heinous self-narrative with a beautiful self-narrative. But the Self-Narrative Account implies that this would not be good for you. On the contrary, it implies that this would bad (or not) for you in the way in which being annihilated would be bad (or not) for you (see Ch. 4, §V).

Winston Smith was forced by the Party's Thought Police to acquire a new implicit understanding of the ongoing course of his life. Smith now implicitly understands his life as a life dedicated to promoting the Party and its ideals. He plans for action in light of this new understanding. And so on. The Thought Police forced Smith to acquire a new, and worse, self-narrative. They thereby harmed Smith, but not in the same way in which they would have harmed Smith had they annihilated him (see Ch. 4, §V). This fact is not consistent with the Self-Narrative Account.

Suppose your child is just old enough to have a self-narrative. You are trying to raise this child so she will eventually become someone with the sort of self-narrative that befits a mature adult. I deny that you are thereby trying to effect the practical equivalent of your child's being annihilated. But this is just what the Self-Narrative Account of survival implies.

Suppose instead that you completely stunted your adult child's psychological growth when that child was seven years old. As a result,

your child never acquired a self-narrative that would befit an adult. Instead, your child has continued, for decades, to have the self-narrative he had when he was seven years old. The Self-Narrative Account implies that you helped your child—when he was seven years old—to survive for the long haul. But I deny that you helped your child by stunting his psychological growth; on the contrary, you grievously harmed him by keeping him from growing up.

The above objections presuppose that seven-year-old children, and Winston Smith, and Nazis, and so on have self-narratives. Some deny that persons in general have self-narratives (§II). This denial might undermine some of this section's objections to the Self-Narrative Account. But this denial will not save that account. On the contrary, this denial implies that that account is false. This is because the Self-Narrative Account of survival is true only if persons in general have self-narratives. In other words—and as we saw in Section II—the Self-Narrative Account of survival implies that persons in general have self-narratives.

Like the Selfer view (Ch. 4, §V), the Self-Narrative Account manages to be both implausibly optimistic and implausibly hopeless at the same time. Here are two examples of the implausible optimism of the Self-Narrative Account: you dread the impossible when you dread, for your mother's sake, her someday living with self-narrative-erasing dementia; and you dread the impossible when you dread, for your own sake, becoming a bad person with a heinous self-narrative. Here are two examples of the implausible hopelessness of that account: you hope for the impossible when you hope, for your child's sake, that that child will have a full and rich adult life, along with a self-narrative that befits such a life; and you hope for the impossible when you hope, for a Nazi's sake, that that Nazi might one day acquire a new, and beautiful, self-narrative.

VI. Other Work for Self-Narrative and the Same Self

Self-narrative might do good philosophical work even if the Self-Narrative Account of survival is false. Here is one example: recognizing the value of having a self-narrative can show us one way in which self-narrative-erasing dementia would be distinctively bad for you.

Another example of the philosophical work self-narrative might do starts with Schechtman's remark:

> According to the narrative self-constitution view, the difference between persons and other individuals (I use the word 'individual' to refer to any sentient creature) lies in how they organize their experience, and hence their lives. (1996, 94)

And Schechtman adds:

> The narrative self-constitution view thus holds firm on the claim that a self-conception sufficiently unlike a traditional linear narrative excludes personhood. (1996, 100)

Thus Schechtman endorses an account of what it is to be a person in terms of having a self-narrative.

So does Charles Taylor. He opens *Sources of the Self* by saying that he wants to explore: 'modern notions of what it is to be a human agent, a person, or a self' (1989, 3). And Taylor (1989, esp. 53–60) takes having a self-narrative to be crucial to being a person, an agent, or a self.

As we saw in Section I, Schechtman takes your self-narrative to create your 'identity'. Similarly, Taylor says:

> My identity is defined by the commitments and identifications which provide the frame or horizon within which I can try to determine from case to case what is good, or valuable, or what ought to be done, or what I endorse or oppose. (1989, 27)

And Taylor adds:

> . . . we saw that our being selves is essentially linked to our sense of the good, and that we achieve selfhood among other selves. Here I have been arguing that the issue of how we are placed in relation to this good is of crucial and inescapable concern for us, that we cannot but strive to give our lives meaning or substance, and that this means that we understand ourselves inescapably in narrative. (1989, 51)

I do not endorse Schechtman's or Taylor's accounts of being a person or of 'identity'. But nothing I have said in this chapter rules out these accounts. In particular, these accounts are perfectly consistent with the falsity of the Self-Narrative Account of survival. Moreover, we get interesting results when we combine these accounts with the falsity of the Self-Narrative Account.

Suppose that being a person is a matter of having a self-narrative. Add that the Self-Narrative Account of survival is false. Then in a case of self-narrative-erasing dementia, a conscious human being who, at a future time, is no longer a person can still have, at that time, what matters in survival for you. So we get the result that you can survive even if you cease to be a person.[1]

Or suppose that your 'identity' is somehow the result of your self-narrative. Add that your current self-narrative might eventually be replaced by another self-narrative. Also add the falsity of the Self-Narrative Account of survival. Then we get the result that you can survive losing one identity and acquiring another. Or suppose that you remain the 'same person' in virtue of having the same self-narrative; then we get the result that you can survive becoming a different person.

Some of the Selfer's claims might also do good philosophical work even if the Selfer view is false. Suppose—as Selfers do—that your values, desires, and projects 'make you the person you are' and give you your 'identity'. Add that this implies that if you were to change with regard to (enough of) those values, desires, and projects, then you would become a 'different person' and would lose your 'identity'. Add that you can survive this sort of change; that is, add that the Selfer view is false (see Ch. 4, §§III–V). Then we get the results that you can survive becoming a different person and also that you can survive losing your identity.

Claims like 'you can survive losing your identity' or 'you can survive becoming a different person' can sound contradictory. But these claims are not contradictory given the above accounts of 'identity' and of 'being the same person'. For given those accounts, these claims amount to the following, which is not contradictory, and

[1] At least, I believe that we get this result. For I say that being numerically identical with is a good answer to:

The Why Question: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

And I think that this is still a good answer (and my defense of that answer is still a good defense) even if we replace the word 'person' with the words 'human being' in the Why Question.

does not even sound contradictory: it can be appropriate for you to first-personally anticipate, and have self-interested concern with regard to, the experiences a person will have at a future time even if that person will not, at that time, have the self-narrative—or the values, desires, and projects—that you now have.

VII. Conclusion

Consider:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

The Self-Narrative Account of survival answers this question in terms of psychological connectedness that is constituted by having the same self-narrative. We have seen that this is not a good answer. But you might still think that the answer to the Why Question should involve narrative. So perhaps you will say that there is a good answer in terms of narrative-constituting psychological continuity. Chapter 6 explores just such an answer, and also explores an answer in terms of agency.

6

Agential Continuity and Narrative Continuity

Consider:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

Chapter 5 focused on an answer to the Why Question in terms of a specific sort of psychological connectedness, namely, having the same self-narrative. This chapter will consider two more answers to the Why Question, one of which—like the answer considered in Chapter 5—involves narrative. But the answers considered in this chapter are in terms of specific sorts of psychological continuity, as opposed to psychological connectedness.

I. The Agential Continuity Account

Christine Korsgaard says:

> It is, I think, significant that writers on personal identity often tell stories about mad surgeons who make changes in our memories or characters. These writers usually emphasize the fact that after the surgical intervention we are altered, we have changed. But surely part of what creates the sense of lost identity is that the person is changed by *intervention*, from outside. The stories might affect us differently if we imagined the changes initiated by the person herself, as a result of her own choice. You are not a different person *just* because you are very different. Authorial psychological connectedness is consistent with drastic changes, provided those changes are the result of actions by the person or reactions for which she is responsible. (1989, 122–3)

Self and Identity. Trenton Merricks, Oxford University Press. © Trenton Merricks 2022.
DOI: 10.1093/oso/9780192843432.003.0007

Korsgaard imagines a person's choosing to act in a way that changes that person's memories and character. Korsgaard is thereby imagining a sort of psychological connectedness, the sort constituted by one's earlier choices resulting in (at least some of) one's later psychological states. Korsgaard calls this sort of psychological connectedness 'authorial psychological connectedness'. But I shall call it 'agential connectedness'.

Korsgaard says:

> You normally think you lead one continuing life because you are one person, but . . . the truth is the reverse. You are one continuing person because you have one life to lead. (1989, 113)

Korsgaard thinks that 'one life' just is a chain of overlapping instances of agential connectedness. This is why she says: 'the choice of any action, no matter how trivial, takes you some way into the future' (1989, 113).

Let *agential continuity* be a chain of overlapping instances of agential connectedness. Korsgaard (1989) thinks that 'you are one continuing person because you have one life to lead', and that you lead 'one life' in virtue of agential continuity. So it is tempting to read Korsgaard (1989) as defending an agential continuity criterion of personal identity over time. In other words, it is tempting to read Korsgaard as defending the claim that agential continuity is necessary and sufficient for (but does not presuppose) personal identity over time. And this reading might seem to be supported by Korsgaard's saying:

> Where I change myself, the sort of continuity needed for personal identity may be preserved, even if I become very different. Where I am changed by wholly external forces it is not. (1989, 123)

But Korsgaard (1989) is not defending an agential continuity criterion of personal identity over time. For she is not defending a criterion of personal identity over time of any sort. Rather, as we shall see, Korsgaard is defending an answer to the Why Question.

Korsgaard says:

> Parfit's conception of the person is recognizably metaphysical in that it is concerned with the theoretical conditions of identity and counting, certainly traditional concerns of metaphysics . . . But our metaphysical concerns about

countability . . . are still just some concerns among others. And they are not obviously the important ones for ethics. (1989, 131n47)

Korsgaard denies that the metaphysics of identity and counting is obviously important for ethics. A proposed criterion of personal identity over time just is a thesis about the metaphysics of identity and counting. So Korsgaard must deny that a criterion of personal identity over time is obviously important for ethics. But Korsgaard (1989) does take having 'one life to lead' to be obviously important for ethics. For example, she says:

Lives conceived of as led by agents may be completely separate even if the unity of those agents is pragmatic rather than metaphysically deep. And if living a life in this sense is what matters, distribution should be over lives, and the agents who lead them. (1989, 129)

Korsgaard also says:

First, the need for identification with some unifying principle or way of choosing is imposed on us by the necessity of making deliberative choices, not by the metaphysical facts. Second, the metaphysical facts do not obviously settle the question: I must still decide whether the consideration that some future person is 'me' has some special normative force for me. It is practical reason that requires me to construct an identity for myself; whether metaphysics is to guide me in this or not is an open question. (1989, 112)

The metaphysical facts obviously settle the metaphysical questions. So surely Korsgaard is not trying to settle this metaphysical question: how must you be related to a person at a future time in order for that person to be identical with you? So Korsgaard is not endorsing a criterion of personal identity over time, not even when she says 'you are one continuing person because you have one life to lead'. Rather, Korsgaard is endorsing a view about when a future person's relation to you 'has some special normative force for' you.

Korsgaard adds:

I have suggested that agents come in different sizes, and that the human body is merely the basic one. If we grant that the unity of agency is a reason for future concern, then we should grant I also have reasons to care for the future of larger agencies of which I am a part. Just as I have a personal concern for my physical future, I may have a *personal* concern for the future of my family, the organization for which I work, a project in which I have been active, or the state of which I am a citizen. (1989, 127)

Korsgaard is not saying: '*L'état, c'est moi.*' That is, she is not claiming that she persists as a future state—is one and the same entity as that state—just in case she enjoys 'unity of agency' with that future state. Rather, Korsgaard is making a claim about her having reasons for 'personal concern' with regard to that state at a future time.

By the same token, Korsgaard is not claiming that you persist as a future person—are one and the same entity as that person—just in case you enjoy 'unity of agency' with that future person. Rather, she is making a claim about your having reasons for 'personal concern' with regard to that person at that future time. So she is making a claim about your being related to that person at that time in a way that has 'some special normative force for' you. So I think that Korsgaard is claiming (among other things) that it is appropriate for you to have future-directed self-interested concern with regard to a person's experiences at a future time just in case you enjoy 'unity of agency' with that person at that time.

You enjoy 'unity of agency' with a person at a future time if and only if you are the same agent as that person at that time. Korsgaard thinks that you are the same agent as a person at a future time if and only if—and because—you and that person lead the same 'one life'. That is, Korsgaard thinks that you are the same agent as a person at a future time if and only if—and because—you are agentially continuous with that person at that time.

So Korsgaard is committed to the view that it is appropriate for you to have future-directed self-interested concern with regard to a person's experiences at a future time if and only if—and because—you are agentially continuous with that person at that time. And if it is appropriate to have future-directed self-interested concern with regard to an experience, then it is appropriate to first-personally anticipate that experience (Ch. 1, §II).

Recall:

> *The What Question*: What is it for a person at a future time to have (at that time) what matters in survival for you?

The answer: its being appropriate for you to first-personally anticipate the experiences that that person will have at that future time; and if that person will have good (or bad) experiences at that time,

its being appropriate for you to have future-directed self-interested concern with regard to those experiences (Ch. 1, §II).

So Korsgaard (1989) is committed to the claim that you survive as a person at a future time if and only if—and because—the way you are now is agentially continuous with the way that person will be at that time. Let this claim be the *Agential Continuity Account of survival*. The Agential Continuity Account implies that being agentially continuous with is a good answer to:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

As Korsgaard says: 'the sort of continuity needed for what matters to me in my own personal identity essentially involves my agency' (1989, 123).[1]

Korsgaard (1989) claims that an answer to (what I call) the Why Question must be motivated by 'practical reasons', and practical reasons alone. This claim is a potential threat to my understanding of how metaphysics is relevant to answering the Why Question (see Ch. 2). So this section will close by showing why I am not persuaded by either of Korsgaard's two ways of defending the claim that an answer to the Why Question must be motivated by practical reasons alone.

Making deliberative choices involves choosing to act on one of your desires as opposed to another. Korsgaard (1989, 111) says: 'To identify with [a] way of choosing [which of your desires to act on] is to be "a law to yourself", and to be unified as such.' So Korsgaard (1989, 112) takes 'the necessity of making deliberative choices' to be a practical reason for regarding yourself as one person now. She adds that you should think of yourself as one person now only for practical reasons. And she takes this addition to imply that only practical reasons may properly motivate an answer to the Why Question. Korsgaard supports this addition with:

[1] The passages quoted in this section all come from Korsgaard (1989). When returning to similar views in her book *Self-Constitution,* Korsgaard replaces the expression 'personal identity' with 'personal or practical identity' (2009, 19–20). This replacement reinforces the point that she is not defending a criterion of personal identity over time.

> Why do you think of yourself as one person now? This problem should seem especially pressing if Parfit has convinced you that you are not unified by a Cartesian Ego which provides a common subject for all your experiences.
> (1989, 109)

But I reply that there is a good reason that is not a practical reason for your thinking of yourself as one person now. That reason is that you are a human animal, a human organism, a living human body (see Ch. 3, §IV). And that human animal is one thing just as much as a 'Cartesian ego' would be one thing.[2] So you are one thing. You are a person now. So you are one person now. And you are (the one thing that is) the common subject of your experiences.

Korsgaard has a second defense of the claim that only practical reasons may motivate an answer to the Why Question. This defense starts with:

> Suppose that Parfit has established that there is no deep sense in which I am identical to the subject of experiences who will occupy my body in the future. In this section I will argue that I nevertheless have reasons for regarding myself as the same rational agent as the one who will occupy my body in the future. These reasons are not metaphysical, but practical. (Korsgaard, 1989, 109)

I think this defense goes wrong right at the start. For I think that persons endure. So there is a 'deep sense' in which you will be identical with 'the subject of experiences who will occupy [your] body in the future'. That 'deep sense' is your being numerically identical with (the person who will be) that subject of experiences. In Parfitian parlance: your being numerically identical with a person at a future time is a 'further fact', distinct from your being psychologically or otherwise connected to or continuous with that person at that time. (See Ch. 2, §III; Ch. 3, §IV.)

[2] Do not conflate *being one thing* and *lacking proper parts*. That would just be a mistake. Moreover, this conflation would not help Korsgaard. On the contrary, it would undermine her practical reason for your regarding yourself as one person now. For even if 'the necessity of making deliberative choices' is a practical reason for regarding yourself as one person now, it is no reason at all for regarding yourself as partless.

II. The Narrative Continuity Account

Alasdair MacIntyre says:

> It is a conceptual commonplace, both for philosophers and ordinary agents, that one and the same segment of human behavior may be correctly characterized in a number of different ways. To the question 'What is he doing?' the answers may with equal truth and appropriateness be 'Digging', 'Gardening', 'Taking exercise', 'Preparing for winter', or 'Pleasing his wife' . . . . We cannot . . . characterize behavior independently of intentions, and we cannot characterize intentions independently of settings which make those intentions intelligible both to agents themselves and to others. (1984, 206)

An agent's actions cannot be characterized apart from that agent's intentions. MacIntyre thinks that an agent's intentions must be understood in a setting, which itself must be part of a narrative. So MacIntyre says:

> Narrative history of a certain kind turns out to be the basic and essential genre for the characterization of human action. (1984, 208)

Chapter 5 considered self-narratives, each of which characterizes an entire life. MacIntyre's remarks suggest a different sort of narrative, the sort that characterizes a single action. So let a *local narrative* be a narrative that characterizes a single action. Add that a local narrative is constituted by the agent's action-characterizing psychological states. So a local narrative might be a story that is constituted by, for example, the reasons that that agent decided to act, what that agent intends to do, that agent's interpretation of what she or he is doing, and that agent's experiences while acting.

Let *local narrative continuity* be a chain of overlapping local narratives. And let the *Narrative Continuity Account of survival* be the following two claims: first, you will survive as a person at a future time if and only if the way you are now is locally narratively continuous with the way that that person will be at that time; and, second, being locally narratively continuous with is a good answer to:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

Because a local narrative is constituted by the agent's action-characterizing psychological states, local narrative continuity is a specific sort of psychological continuity. So the Narrative Continuity Account's answer to the Why Question is in terms of a specific sort of psychological continuity. So that account's answer is the kind of answer I want to explore (see Ch. 3, §V).[3]

Jeanine Weekes Schroer and Robert Schroer say:

> Under our account, the narrative that frames Larry's experience of making coffee via his childhood memories of family and feelings of security need not be interlaced or otherwise connected to the narrative explanations that frame and explain some of his other experiences/actions. The narrative explanation that frames his experience of balancing his checkbook, for instance, may have no connection to his childhood memories of family or feelings of security. To put the point more generally, the requirement that the subject interpret the events of his or her life via 'explanations that mirror the logic found in stories' does not entail that those narrative explanations all fit together into a single, overarching story of the person's life. (2014, 458)

Schroer and Schroer are not focused on narratives that tell the 'single, overarching story of the person's life'. So they are not focused on self-

[3] The Narrative Continuity Account is not supposed to be the Self-Narrative Account under a different label. So let us explicitly add that a local narrative characterizes less than an entire life. Then a local narrative is not a self-narrative. So the Narrative Continuity Account is a genuine alternative to the Self-Narrative Account.

MacIntyre's remarks quoted in the text really do suggest the idea of a local narrative, which is why I quoted them. But MacIntyre himself seems to deny that there are local narratives. For MacIntyre seems to take the narratives that render actions intelligible to characterize an entire life, and even more than an entire life. MacIntyre says:

> . . . in successfully identifying and understanding what someone else is doing we always move towards placing a particular episode in the context of a set of narrative histories, histories both of the individual concerned and of the settings in which they act and suffer. (1984, 211)

Charles Taylor also seems to reject the idea of a local narrative, at least with regard to characterizing non-trivial actions. He seems to think that characterizing a non-trivial action requires a self-narrative. Taylor says:

> Thus making sense of my present action, when we are not dealing with such trivial questions as where I shall go in the next five minutes but with the issue of my place relative to the good, requires a narrative understanding of my life, a sense of what I have become which can only be given in a story. (1989, 48)

narratives. Rather, they are focused on just enough narrative to characterize 'experiences/actions', such as making coffee or balancing a checkbook. And the narratives that Schroer and Schroer take to characterize an agent's action seem to be constituted by that agent's own psychological states: 'memories' and 'feelings' and how the agent 'interpret[s] the events of his or her life'. So Schroer and Schroer are focused on local narratives.

Schroer and Schroer (2014) defend the view that personal identity over time is reduced to unbranching local narrative continuity holding between 'person stages', that is, the temporal parts of perduring persons. In reaction to the possibility of branching local narrative continuity, they say:

> Another well-known Reductionist response to cases of fission comes from Parfit (1984), who argues that fission cases reveal that, when it comes to survival, 'identity doesn't matter'. Although cases of fission destroy identity, what 'really matters' with regard to survival—relations of [generic] psychological continuity between person stages—continues to exist. Again, we see no reason why the same reply cannot be given by Reductionists, like us, who emphasize narrative continuity over [Parfit's generic] psychological continuity. Although cases of fission destroy personal identity, what 'really matters' with regard to survival—narrative continuity between person stages—continues to exist. (2014, 467)

Thus Schroer and Schroer (2014) take the Narrative Continuity Account of survival to be a live option.

The Narrative Continuity Account answers the Why Question in terms of overlapping local narratives. Local narratives characterize actions. Only agents act. So the Narrative Continuity Account takes agency to be central to an answer to the Why Question. This makes the Narrative Continuity Account somewhat like the Agential Continuity Account. And the Narrative Continuity Account and the Agential Continuity Account seem to be alike in another way. They seem to deliver the same 'verdict' with regard to whether a particular future person will have what now matters in survival for you.

To begin to see why I say this, suppose that you are related to a person at a future time by local narrative continuity. Then you are related to that person at that future time by overlapping instances of action-characterizing psychological connectedness. Then it seems that

you are also related to that person at that future time by overlapping instances of choosing to act that result in the having of various psychological states. So it seems that you are agentially continuous with that person at that future time. So if you will survive as a person at a future time by the lights of the Narrative Continuity Account of survival, then it seems that you will survive as that person at that time by the lights of the Agential Continuity Account of survival.

I think that defenders of the Narrative Continuity Account should assume that every action is characterized by a local narrative. Suppose they are right. And suppose that you are agentially continuous with a person at a future time. Then you are related to that person at that future time by overlapping instances of choosing to act that result in the having of various psychological states. Then it seems that—because every action is characterized by a local narrative—you are also related to that person at that future time by overlapping instances of action-characterizing psychological connectedness. So it seems that you are related to that person at that future time by local narrative continuity. So if you will survive as a person at a future time by the lights of the Agential Continuity Account of survival, then it seems that you will survive as that person at that time by the lights of the Narrative Continuity Account of survival.

III. Agential Continuity, Narrative Continuity, and Numerical Identity

Consider:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

The Agential Continuity Account of survival answers this question in terms of agential continuity. The Narrative Continuity Account of survival answers this question in terms of local narrative continuity. My answer is not in terms of either of these sorts of continuity, but is instead in terms of numerical identity. This is my first reason for taking both the Agential Continuity Account and also the Narrative Continuity Account to compete with my answer to the Why Question.

I think that you are a human organism (Ch. 3, §IV). And I think that human organisms can endure in the absence of both agential continuity and also local narrative continuity. So I conclude that you can be numerically identical with a (conscious) person at a future time even if you are not agentially continuous with, or locally narratively continuous with, that person at that time. So I take my answer to the Why Question to imply that you can survive as a person at a future time even if you are not agentially continuous with, or locally narratively continuous with, that person at that time. But this implication is false if either the Agential Continuity Account or the Narrative Continuity Account is true. This is my second reason for taking both of those accounts to compete with my answer to the Why Question.

Suppose that persons endure. Suppose also that the criterion of personal identity over time is agential continuity. Then agential continuity might be a good answer to the Why Question because numerical identity is also a good answer to the Why Question (see Ch. 2, §III). Then we should reject my first reason for taking the Agential Continuity Account to compete with my answer to the Why Question. And if the criterion of personal identity over time is local narrative continuity, we should reject my first reason for taking the Narrative Continuity Account to compete with my answer to the Why Question.

Again, suppose that the criterion of personal identity over time is agential continuity. Then you cannot be numerically identical with a person at a future time but not agentially continuous with that person at that time. Then we should reject my second reason for taking the Agential Continuity Account to compete with my answer to the Why Question. And if the criterion of personal identity over time is local narrative continuity, we should reject my second reason for taking the Narrative Continuity Account to compete with my answer to the Why Question.

Suppose that agential continuity is the criterion of personal identity over time. Then we should do more than reject my two reasons for taking the Agential Continuity Account to compete with my answer to the Why Question. We should also conclude that the Agential Continuity Account does not compete with my answer. And if local narrative continuity is the criterion of personal identity over time, we

should conclude that the Narrative Continuity Account does not compete with my answer.

It seems that you will be agentially continuous with a person at a future time if and only if you will be locally narratively continuous with that person at that time (§II). Add that one or the other of these two sorts of continuity is the criterion of personal identity over time. Then your being related to a person at a future time by either sort of continuity is—if you endure—both necessary and sufficient for your being numerically identical with that person at that time. So it seems that if either sort of continuity is the criterion of personal identity over time, neither the Agential Continuity Account nor the Narrative Continuity Account competes with my answer to the Why Question.

Let us assume—as I myself believe—that neither agential continuity nor local narrative continuity is the criterion of personal identity over time.[4] Then the Agential Continuity Account and the Narrative Continuity Account compete with my answer to the Why Question, and do so for the two reasons presented at the start of this section. Then I want to keep discussing the Agential Continuity Account and the Narrative Continuity Account. For I want to explore answers to the Why Question that both compete with my answer, and also are in terms of—or at least imply—psychological connectedness or psychological continuity (see Ch. 3, §V).

The Agential Continuity Account and the Narrative Continuity Account compete with my answer to the Why Question. So if my answer to the Why Question is a good answer, then both of those accounts are false. So my defense of my answer to the Why Question is thereby a defense of the falsity of those accounts (Ch. 2, §I). So I conclude that each of those accounts is false. But you might draw a different conclusion. You might endorse one of those accounts. If one of those account is true, then my answer to the Why Question is not a good answer. So you might conclude that my answer is not a good

[4] One defense of this assumption starts with the claim that each of us is an organism. Here is another defense: agential continuity and narrative continuity can branch; since we endure, personal identity over time cannot branch (Ch. 3, §IV); so neither agential continuity nor local narrative continuity is the criterion of personal identity over time.

answer, and that something or other must have gone wrong in my defense of that answer.

I want to respond to this threat to my answer and to my defense of that answer. So Section IV will present reasons to conclude that both the Agential Continuity Account and also the Narrative Continuity Account are false, reasons that do not presuppose my answer to the Why Question, or my defense of that answer. And if those accounts are false, their competing with my answer is no reason to reject my answer, or to think that something or other must have gone wrong in my defense of that answer.

IV. Some Significant Transformations

Consider the following two cases:

> Case One: A vampire is about to bite you. The vampire pauses—still, you never see it coming—and then bites. As a result, you are transformed into a vampire. The way you were immediately before transformation is not agentially continuous with the way you are upon transformation. But you would have chosen to become a vampire, had you been given the choice.
>
> Case Two: A vampire is about to bite you. The vampire pauses. You sense the pause, and assume the vampire is having second thoughts. So you shove your neck onto the vampire's teeth, intending to become a vampire. You succeed and are transformed into a vampire. The way you were immediately before transformation is agentially continuous with the way you are upon transformation. (This case is otherwise as much like Case One as possible.)

In Case One, your becoming a vampire is not a result of your own action. This is why Case One can include the claim that the way you were immediately before transformation is not agentially continuous with the way you are upon transformation. And because Case One includes this claim, the Agential Continuity Account of survival implies that you do not survive becoming a vampire in Case One.[5]

[5] Case One presupposes that you persist as a vampire. If agential continuity is the criterion of personal identity over time, this presupposition is false. So if you think that agential continuity is the criterion of personal identity over time,

In Case Two, your becoming a vampire is the result of your own action, specifically, your shoving your neck onto a vampire's teeth with the intention of becoming a vampire yourself. So—as Case Two explicitly notes—the way you were immediately before transformation is agentially continuous with the way you are upon transformation. So the Agential Continuity Account of survival implies that you survive becoming a vampire in Case Two. As Korsgaard says:

> Where I change myself, the sort of continuity needed for personal identity may be preserved, even if I become very different. Where I am changed by wholly external forces it is not. This is because the sort of continuity needed for what matters to me in my own personal identity essentially involves my agency.
> (1989, 123)

The Agential Continuity Account implies the following conjunction: you do not survive becoming a vampire in Case One and you do survive becoming a vampire in Case Two. If this conjunction is false, then the Agential Continuity Account is false. And this conjunction is false.

One reason that I say that this conjunction is false is that I think that you survive becoming a vampire in both cases. I think this because of my answer to the Why Question and my belief that you endure. But set this aside. For in this section I want to object to the Agential Continuity Account (and the Narrative Continuity Account) without relying on my answer to the Why Question (see §III).

You might say that you do not survive becoming a vampire in either Case One or Case Two. Or you might say that you do survive becoming a vampire in both cases. Or you might say that you do not know what to say. But the one thing that you should not say is that you survive becoming a vampire in one of these cases but not in the other. For these cases differ only in what causes some teeth to enter your neck. And I think that it is false that this difference would make all the difference in whether you survive.

you should object to Case One, and to other moves made in this section. But if you think that agential continuity is the criterion of personal identity over time, you should not take the Agential Continuity Account to compete with my answer to the Why Question (§III). So you and I agree on what is most important here: the Agential Continuity Account is no reason to reject my answer, or to think that something or other must have gone wrong in my defense of that answer.

Defenders of the Agential Continuity Account might object that I have overlooked another way in which Case One differs from Case Two. They might object that those cases differ with respect to whether agential continuity relates you-before-teeth-enter-neck to you-after-transformation. But I reply that this difference in agential continuity just is the difference already pointed out, a difference in what causes some teeth to enter your neck.

It might help to contrast numerical identity holding over time with agential continuity holding over time. Numerical identity holding over time is a 'further fact' distinct from any (alleged) criterion of personal identity holding over time (Ch. 2, §III). But agential continuity is not a 'further fact' distinct from overlapping instances of agential connectedness. Agential continuity just is overlapping instances of agential connectedness. This is why the difference in agential continuity between Cases One and Two just is the difference already pointed out, a difference in what causes some teeth to enter your neck.

Again, the Agential Continuity Account of survival implies that a difference in what causes some teeth to enter your neck can make all the difference in whether you survive or do not survive becoming a vampire. But I deny that whether it is appropriate for you to first-personally anticipate, and extend your future-directed self-interested concern to, the experiences a vampire will have at a future time could turn on whether you shove your neck onto some teeth or instead those teeth are shoved into your neck. So the following conjunction is false: you do not survive becoming a vampire in Case One and you do survive becoming a vampire in Case Two. So the Agential Continuity Account is false.

Consider:

> Case Three: Vampires are incredibly evil. You find the thought of becoming a vampire abhorrent and terrifying. You definitely do not want to become a vampire. But vampires are crafty. And—you will never see it coming—a vampire will soon bite you. As a result, you will be transformed into a vampire. The way you are now is not agentially continuous with the way you will be upon transformation.

Your being an evil vampire at a future time will be bad for you—for the you of right now—only if you will survive as that vampire (see Ch. 1, §§II–III). The Agential Continuity Account of survival implies that

you will not survive becoming an evil vampire in Case Three. So that account implies that in Case Three, your being an evil vampire at a future time will not be bad for you. Rather, that account implies that the events of Case Three will be bad (or not) for you only in the way in which being annihilated would be bad (or not) for you.

Maybe being annihilated would not be bad for you. But being transformed into an evil vampire against your will would be bad for you. Maybe being annihilated would be bad for you only in that it would deprive you of the goods that come with continued survival. But it is false that being transformed into an evil vampire would be bad for you only in that it would deprive you of the goods that come with continued survival. More generally, being transformed into an evil vampire, against your will, would be bad for you, and not only in the way in which being annihilated would be bad (or not) for you (cf. Ch. 4, §V).

Imagine this scene from a vampire movie. You have just been bitten by an evil vampire. Your transformation has not yet started. But it soon will. You would rather cease to exist than become an evil vampire. So you ask your friends to kill you, and to do so for your sake. They balk, so you beg. This scene makes sense. But not given the Agential Continuity Account. For given that account, and as far as self-interested concern goes, you should be indifferent between ceasing to exist and becoming an evil vampire against your will. But you are not indifferent. Hence the begging. I conclude that the Agential Continuity Account is false.

Consider:

> Case Four: You will soon be transformed 'by wholly external forces' into a much happier, smarter, wiser, and more loving person. So the way you are now is not agentially continuous with the way you will be after transformation.

The Agential Continuity Account implies that the events of Case Four would be bad (or not) for you in the same way—and to the same degree—that the events of Case Three would be bad (or not) for you. For the Agential Continuity Account implies that both cases would be bad (or not) for you in the way in which being annihilated would be bad (or not) for you.

But I deny that Case Three and Case Four are both bad (or not) for you in the same way and to the same degree. It is worse for you to be transformed into an evil vampire who will not be agentially continuous with you than it is for you to be transformed into a much happier, smarter, wiser, and more loving person who will not be agentially continuous with you. Again, the Agential Continuity Account is false.

Recall that Winston Smith has values, desires, projects, and a self-narrative that are not approved of by the Party. The Party's Thought Police take note, and forcibly 'reintegrate' Smith. That is, they turn Smith, by wholly external forces, into the sort of person approved of by the Party. I think that the Thought Police have harmed Smith, but not in the way that they would have harmed Smith had they annihilated him. This is another reason to conclude that the Agential Continuity Account is false (cf. Ch. 4, §V; Ch. 5, §V).

I learn that the Thought Police are coming for me. So I am afraid. One of the things I am afraid of is being turned, against my will, into the sort of person approved of by the Party. I fear this in a way that differs from how I would fear being annihilated. And I fear this in a way that differs from how I would fear your being turned into the sort of person approved of by the Party. For the way in which I fear this involves my first-personally anticipating, and having future-directed self-interested concern with regard to, the experiences of someone who will be the way the Party wants people to be. I think my fear is appropriate. But the Agential Continuity Account implies that it is not. So I conclude, again, that that account is false.

Perhaps it is possible that the way you are now is agentially continuous with the way you were when you were a child. So perhaps the Agential Continuity Account of survival does not automatically imply that no child survives becoming an adult. But going through puberty is like being bitten by a vampire. Growing up happens to a child. (And if given a choice, some children would choose not to grow up.) So I think that the way some children are now is not agentially continuous with the way that they will be when adults. So the Agential Continuity Account has the result that those children will not survive growing up (cf. Ch. 4, §IV; Ch. 5, §IV). More generally, that account has the result that no one can survive a significant transformation caused entirely by circumstances beyond his or her control. But this result is

false. You can—and probably have—survived such a transformation. None of us is immune to 'wholly external forces'.

Each of this section's objections to the Agential Continuity Account of survival turns, in one way or another, on the idea that the Agential Continuity Account has a false result regarding whether someone would survive in this or that scenario. When it comes to whether someone would survive in this or that scenario, the Narrative Continuity Account seems to deliver the same 'verdict' as does the Agential Continuity Account (§II). So I conclude that each of this section's objections to the Agential Continuity Account is also an objection to the Narrative Continuity Account.

V. Other Work for Agential Continuity and Narrative Continuity

Suppose that you are the 'same agent' as a person at a future time just in case you are agentially continuous with (and locally narratively continuous with) that person at that time. Add that agential continuity is not necessary for enduring and that numerical identity delivers survival (Ch. 2, §I). Then conclude that it is possible for a person at a future time to have, at that time, what matters in survival for you even if that person will not be the same agent as you.

It is possible that you will someday cease to regard yourself as the same agent as a future person. Korsgaard (1989, 113–14) says that regarding yourself as the same agent as a future person is part of what constitutes your being an agent. Suppose she is right. Then it is possible that you will someday not be an agent. Add that you endure and that numerical identity delivers survival. Then conclude that it is possible that a person at a future time will have, at that time, what matters in survival for you even if that person will not be an agent.

Claims like 'you can survive as someone who is not the same agent as you' or 'you can survive but not be an agent' can sound contradictory. But these claims are not contradictory given the above account of being the 'same agent' and given Korsgaard's view about what constitutes being an agent. For example, 'you can survive as someone who is not the same agent as you' is not contradictory if it amounts to: it can be appropriate for you to first-personally anticipate, and have future-directed self-interested concern with regard to, the experiences a person will have at a future time even if the way you are now is not

agentially (or locally narratively) continuous with the way that person will be at that time (cf. Ch. 5, §VI).

I do not endorse the above account of being the 'same agent'. Nor do I endorse Korsgaard's view of what constitutes being an agent. But nothing I have said in this chapter rules out that account or that view. In particular, that account and that view are perfectly consistent with the falsity of the Agential Continuity Account of survival and with the fasity of the Narrative Continuity Account of survival.

VI. More on Psychological Connectedness and Psychological Continuity

Recall:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

My answer is your being numerically identical with that person at that time. My argument for that answer requires that persons persist by enduring (Ch. 2, §I).

Chapter 2 (§II) showed that perdurance's signature way of rendering change non-contradictory implies that you have no determinate mental properties and none (or few) of the determinate physical properties that you think you have. I believe that this is a serious problem for perdurance. And I suspect that few will be attracted to stage theory's idea that each of us is an instantaneous stage. So I think that endurance looks pretty good. But I do not have more to say to those who oppose my answer to the Why Question because they deny that persons endure.

I do have more to say to those who oppose my answer to the Why Question because they claim that every good answer must be in terms of—or at least imply—some sort of psychological connectedness or psychological continuity. I say: that claim is false. More carefully, that claim is either false, or instead no reason to oppose my answer to the Why Question. Or so I shall argue for the rest of this section. This section's arguments build on arguments from Chapters 4 and 5 and from the preceding sections of this chapter.

Suppose, for reductio, that every good answer to the Why Question must be in terms of—or imply—psychological connectedness. Then we cannot survive transformations that eventually result in a break in psychological connectedness. But we can survive some such transformations. For example, a small child can survive transforming into an adult (Ch. 4, §IV; Ch. 5, §IV). So it is false that every good answer to the Why Question must be in terms of—or imply—psychological connectedness.[6]

Some transformations that eventually result in a break in psychological connectedness would be bad for you, but not in the way that annihilation would be bad for you. And other transformations that eventually result in a break in psychological connectedness could be good for you. All this implies—as we saw in Chapters 4 (§V) and 5 (§V)—that you can survive these transformations. So I conclude, again, that it is false that every good answer to the Why Question must be in terms of—or imply—psychological connectedness.

This conclusion is important because the answers most commonly given to the Why Question that compete with my answer are in terms of—or imply—psychological connectedness. This goes back at least to Leibniz, who thinks that the answer to the Why Question must imply being connected by memory (Ch. 1, §§I–II). Moreover, the currently predominant view seems to be that the answer to the Why Question must imply being psychologically connected in virtue of being alike with regard to values, desires, and projects (Ch. 4, §II). And then there is Parfit:

> I said earlier that what matters in survival could be provisionally referred to as 'psychological continuity'. I must now distinguish this relation from another, which I shall call 'psychological connectedness' . . . 'Psychological connectedness' . . . requires the holding of these direct psychological relations . . . . 'Psychological continuity' . . . only requires overlapping chains of direct psychological relations . . . . Now that we have distinguished the general relations of psychological continuity and psychological connectedness, I suggest that connectedness is a more important element in survival.
>
> (1971, 20–1, quoted in Ch. 3, §II; see also Parfit, 1984, 262)

[6] The sort of psychological connectedness here is supposed to be substantive (cf. Ch. 4, §I; Ch. 5, §§I–II). So take claims about psychological connectedness here to be claims about substantive psychological connectedness.

My answer to the Why Question is not in terms of—and does not imply—psychological connectedness. But this is no reason to object to my answer. For, as we have seen, it is false that every good answer to the Why Question must be in terms of—or even imply—psychological connectedness.

Suppose that psychological continuity is a good answer to the Why Question. But also suppose—for the sake of argument—that psychological continuity is the criterion of personal identity over time. Then psychological continuity and personal identity go hand in hand. Then we should conclude that numerical identity is also a good answer to the Why Question, at least if persons endure. So we should deny that every good answer to the Why Question must be in terms of psychological continuity (cf. §III and Ch. 2, §III).

Suppose—for the sake of argument—that psychological continuity is necessary for personal identity over time. In other words, suppose that personal identity over time implies psychological continuity. Then your being numerically identical with a person at a future time implies that you are psychologically continuous with that person at that time. Then my answer to the Why Question is consistent with the claim that every good answer to the Why Question must imply psychological continuity. So the claim that every good answer to the Why Question must be in terms of—or at least imply—psychological continuity is no threat to my answer (cf. §III).

Suppose—as I myself believe—that psychological continuity is not the criterion of personal identity over time and, moreover, is not even necessary for personal identity over time (see Ch. 3, §IV). Then my answer to the Why Question is not consistent with the claim that every good answer to the Why Question must be in terms of—or at least imply—psychological continuity. Then that claim competes with my answer. But then—so I shall now argue—that claim is false.

The Party's Thought Police forcibly 'reintegrated' Smith. They turned Smith into the sort of person approved of by the Party. Their particular way of achieving reintegration involved a break in Smith's psychological continuity. But Smith persisted through this break. (Recall that we are supposing that psychological continuity is not necessary for personal identity over time.) So Smith himself became the sort of person approved of by the Party. I think that the Thought Police harmed Smith, but not in the way in which they would have

harmed Smith had they annihilated him. So it is false that every good answer to the Why Question must be in terms of—or imply—psychological continuity (cf. Ch. 4, §V; Ch. 5, §V; §IV).

I learn that I am next. The Thought Police are going to reintegrate me via a process that involves a break in psychological continuity. So I am afraid. One of the things I am afraid of is my becoming the sort of person approved of by the Party. I fear this in a way that differs from how I would fear being annihilated. And I fear this in a way that differs from how I would fear your becoming the sort of person approved of by the Party. For my fear includes first-personally anticipating, and having future-directed self-interested concern with regard to, the experiences of someone who will be the way the Party wants people to be. I think my fear is appropriate. So I conclude that it is false that every good answer to the Why Question must be in terms of—or imply—psychological continuity (cf. Ch. 4, §V; Ch. 5, §V; §IV).

Imagine that a new genetic test has revealed that I shall someday suffer an abrupt and total break in psychological connectedness, and so a break in psychological continuity. I think that this will be bad for me, that is, for the me of right now. Moreover, I do not think that this will be bad for me in the way that my someday being annihilated would be bad (or not) for me. So I conclude that I shall survive this break in psychological continuity (see Ch. 1, §§II–III). So it is false that every good answer to the Why Question must be in terms of—or imply—psychological continuity (cf. §IV).

Finally, pretend that every good answer to the Why Question must be in terms of—or at least imply—psychological continuity. Then future-directed self-interested concern would be rendered appropriate only by psychological continuity, or by something that implies psychological continuity. I assume that the way you are now is not psychologically continuous with the way you will be when permanently unconscious and comatose. So—given what we are pretending—it is not appropriate for you to have future-directed self-interested concern with regard to something bad that will happen to you while you are permanently unconscious and comatose. But I think this is appropriate (see Ch. 1, §III). This is another reason that I deny that every good answer to the Why Question must be in terms of—or imply—psychological continuity.

My answer to the Why Question is not in terms of psychological connectedness or psychological continuity. My answer does not imply psychological connectedness. And my answer might not imply psychological continuity. So some will object to my answer on the grounds that every good answer must be in terms of—or at least imply—psychological connectedness or psychological continuity. But we have seen that this objection is mistaken. For the claim that every good answer to the Why Question must be in terms of—or imply—psychological connectedness or psychological continuity is either false, or instead no threat to my answer.

VII. Conclusion

Consider:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

My answer is your being numerically identical with that person at that time.

I endorse my answer because of the argument given in Chapter 2 (§I). I do not endorse my answer because I have ruled out every possible competing answer and my answer is the last one standing. For I have not ruled out every possible competing answer. I have not even examined every possible competing answer. I bet there are a lot.

But I have examined—and tried to rule out—what I take to be the far and away most tempting kind of competing answer. For I have argued for the conclusion that no one should endorse an answer to the Why Question that both competes with my answer and also is in terms of—or implies—psychological connectedness or psychological continuity. My argument for this conclusion runs from Chapter 3 (§V) through Section VI above.

Chapter 7 will not oppose new competitors to my answer to the Why Question. Nor will it offer a new defense of my answer. Nor will it offer a new defense of my claim that every good answer must imply numerical identity (Ch. 3, §I). Instead, Chapter 7 puts the claims defended thus far in this book through their paces by applying them to a new topic: personal immortality.

7

The Hope of Glory

Leibniz's King of China, Parfit's double brain hemisphere transplant, and Paul's vampires have all played a role in this book's exploration of what matters in survival. And you could read this chapter as continuing to explore what matters in survival with one last thought experiment, a thought experiment involving a certain sort of personal immortality. Or you could read this chapter as applying the lessons of the preceding chapters to the philosophy of religion.

I begin this chapter by using the distinction between persistence and survival to clarify the idea of personal immortality. I then show how claims defended in earlier chapters allow us to respond to familiar objections to the possibility and to the desirability of personal immortality. To give just one example, this chapter's response to the 'Tedium Objection' to immortality builds on Chapter 4's conclusion that the Selfer view is false.

I. The Hope of Survival

You will die. Suppose that you will later live again and be 'glorified'. Suppose further that you will then live forever with God and others. And add that all this will be good for you, that is, for the you of right now. Borrowing a phrase from the Apostle Paul, let us call the hope for this sort of personal immortality 'the hope of glory'.

To be glorified is to be profoundly transformed. I do not know the details. But let us assume that this transformation includes becoming perfectly loving. This transformation might also include becoming

Self and Identity. Trenton Merricks, Oxford University Press. © Trenton Merricks 2022.
DOI: 10.1093/oso/9780192843432.003.0008

more creative, or smarter, or funnier. It might even include—hat tip to Origen—becoming spherical.[1]

A future experience will be good for you—that is, for the you of right now—only if the person who will have that experience at a future time will also have, at that time, what matters in survival for you (Ch. 1, §§II–III). The hope of glory is supposed to be a hope for something that will be good for you. So the hope of glory includes the hope of surviving forever, as opposed to merely persisting forever.

The hope of glory includes the hope of surviving forever. Something similar goes for personal immortality more generally. To see why I say this, pretend that there will be an endless afterlife in which you persist, but do not survive. So it is not appropriate for you to first-personally anticipate, or have future-directed self-interested concern with regard to, any experiences you will have in this afterlife (Ch. 1, §II). Nor will any good (or bad) experiences you will have in this afterlife be good (or bad) for you, that is, for the you of right now (Ch. 1, §II). I do not think that this sort of afterlife delivers bona fide personal immortality. Rather, this sort of afterlife delivers the practical equivalent of annihilation. So I conclude that the idea of personal immortality is not the idea of endless persistence, but is instead the idea of endless survival.

I am not alone. Recall Leibniz's remark from Chapter 1 (§I):

> . . . the immortality required in morality and religion does not consist merely in this perpetual subsistence common to all substances. (1989, 66)

And here is William James:

> . . . the Soul, however, when closely scrutinized, guarantees no immortality of a sort *we care for*. The enjoyment of the atom-like simplicity of their substance *in saecula saeculorum* would not to most people seem a consummation devoutly to be wished. (1890, 348)

I take Leibniz and James to be claiming that bona fide personal immortality is not the same thing as merely endless persistence. And I agree. Bona fide personal immortality is, instead, the same

[1] The Second Council of Constantinople condemned Origen for teaching that our resurrected bodies will be spherical. Alas, Origen did not really teach this; see Chadwick (1948) and Ramelli (2019, 172).

thing as there always being someone who will have what matters in survival for you.

I do not agree with everything that Leibniz and James say. In particular, we disagree about how to answer:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

My answer is in terms of numerical identity (Ch. 2, §I). Leibniz's answer seems to be in terms of persistence plus memory (Ch. 1, §§I–II). And James's answer seems to be in terms of persistence plus psychological continuity that is constituted by conscious states. For James says:

> The substance must give rise to a stream of consciousness continuous with the present stream, in order to arouse our hope, but of this the mere persistence of the substance *per se* offers no guarantee. (1890, 348)

I said above that personal immortality is the same thing as there always being someone who will have what matters in survival for you. And I shall proceed as if personal immortality is the same thing as, at some point after you die, there always being someone who will have what matters in survival for you. For this account of personal immortality is good enough for our purposes. In particular, this account makes it clear that personal immortality is not the same thing as endless persistence.

But I do not think that this account of personal immortality is exactly right. For example—and contrary to this account—I think that your enjoying personal immortality is consistent with there being a brief moment a million years from now when no one will have what matters in survival for you. But I shall not try to get this account exactly right. For doing so would raise difficult questions that are not relevant to the points defended in this chapter. Here is just one such question: can you have personal immortality if you will be in an unconscious coma—or even fail to exist—for one decade every twenty years, ad infinitum?

I answer the Why Question in terms of numerical identity. So I take numerical identity (and being conscious) to be sufficient for survival (Ch., 1, §III; Ch. 2, §I). I think that every good answer to the Why Question implies numerical identity. So I take numerical identity to be

necessary for survival (Ch. 3, §I). So I conclude that you will enjoy personal immortality if and only if, at some point after you die, there will always be someone (conscious) who is numerically identical with you. I myself believe that the hope of glory will be satisfied, and for us all. So I conclude that at some point after you die there will always be someone who is numerically identical with you.

Again, you will enjoy personal immortality if and only if there will always be someone (conscious) who is numerically identical with you. This is consistent with denying that personal immortality is the same thing as there always being someone (conscious) who is numerically identical with you. In particular, this is consistent with my claim that personal immortality is, instead, the same thing as there always being someone who will have what matters in survival for you. That is, personal immortality is the same thing as, at some point after you die, there always being a person whose experiences it is appropriate for you to first-personally anticipate and with regard to whose good (or bad) experiences it is appropriate for you to have future-directed self-interested concern (cf. Ch. 1, §II–III).

The main points defended in this chapter are independent of my own particular view about what it would take for there always to be someone numerically identical with you. But I do think that you are a human organism, a human animal, a living human body (Ch. 3, §IV). And, for what it is worth, I think that there will always be someone numerically identical with you just in case your body (= you) will always exist at some point after your death. That is, your body (= you) will be resurrected and then live forever.[2]

II. The Hope of Transformation

When you were a small child, you did not have many of your current values, desires, and projects, or your current self-narrative. So I think

[2] Chapter 3 (§IV) gives one argument for the claim that you are a human organism. Another argument for that claim turns on the idea that our hope for eternal life goes hand in hand with our hope for the resurrection of the body (see Merricks, 2009). How can the resurrected body be numerically identical with the body that died, and (let us add) was cremated and ceased to exist? For my answer, see Merricks, 2001b.

that when you are glorified and a billion years old, you will no longer have many of your current values, desires, and projects, or your current self-narrative. I think this for two reasons. First, glorification will presumably be as big a transformation as are the transformations that often occur in this life, transformations such as learning to read, going through puberty, and becoming a mother. Second, a billion years dwarfs the amount of time that has passed since your childhood.

Again, when you are glorified and a billion years old, you will no longer have many of your current values, desires, and projects, or your current self-narrative. So both the Selfer view and also the Self-Narrative Account of survival imply that you could not survive becoming glorified and persisting for a billion years, much less forever. So the Selfer view and the Self-Narrative Account imply that the hope of glory is the hope for the impossible. But the Selfer view and the Self-Narrative Account are both false (Chs 4 and 5). So they give us no good reason to conclude that the hope of glory is the hope for the impossible.

Gregory of Nyssa takes glorification to be a process of endless transformations, transformations that make us more and more like God (see, e.g. Smith, 2018). With this in mind, consider the following from Jeff McMahan:

> . . . one might imagine the prospect of becoming like a god. Imagine the possibility of becoming vastly more intelligent and developing a vastly richer and deeper range of emotions, including emotions of which one cannot now form any conception. One would be as different from oneself now, in terms of psychological capacities, as one is now from a dog (or, more to the point, as different from oneself now as a dog would be from itself if it were to become a person). One would be, in short, so utterly psychologically remote from oneself as one is now that one may now have little or no egoistic reason to want to become that way. Even if the transformation would be identity-preserving and would lead to a state that would be clearly superior to one's present state, it would be too much like becoming someone else—and, of course, losing oneself in the process—to be very desirable from an egoistic point of view.
> (2002, 321–2)

Suppose that glorification involves our acquiring new psychological capacities, capacities that differ greatly from the ones we now enjoy, capacities of which we cannot now form any conception. Then McMahan's remarks suggest that even if glorification 'would be

identity preserving and would lead to a state that would be clearly superior', glorification would not be 'very desirable from an egoistic point of view'. So McMahan's remarks suggest that it would not be appropriate for you to have future-directed self-interested concern with regard to the experiences that you will have after you have been glorified. So you would not survive becoming glorified. So the hope of glory is the hope for the impossible.

I have two replies. The first reply begins by noting that a small child's becoming a wise and mature and intelligent adult is a process that will involve that child's acquiring new psychological capacities, at least some of which that child cannot now form any conception. A small child can survive growing up (see esp. Ch. 4, §IV). So I conclude that we can survive transformations that involve acquiring new psychological capacities, even capacities of which we cannot now form any conception. So we can survive becoming glorified even if this involves our acquiring new psychological capacities, even capacities of which we cannot now form any conception.

My second reply turns on my answer to:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

My answer is in terms of numerical identity (Ch. 2, §I). So enduring (conscious) persons survive all 'identity-preserving' transformations. This includes identity-preserving transformations with regard to psychological capacities. So we can survive becoming glorified even if this involves psychological transformations of the sort McMahan—and Gregory of Nyssa—have in mind, at least if those transformations are identity preserving.

And I think that those transformations can be identity preserving. For example, I think that each of us can be numerically identical with someone who will be perfectly loving, at least eventually, at least if we are given enough time. More generally, I think that each of us can be numerically identical with a glorified person who will exist in a billion years, despite the fact that the way such a person will be differs greatly in psychological and other ways from the way each of us is now. I think this partly because we already know that human persons persist

through astonishing amounts of psychological and other sorts of change, such as, for example, the changes that occur as someone goes from small child to young adult to middle-aged academic to elderly Nobel Laureate.

III. The Hope and Psychological Continuity

I think that the hope of glory includes psychological transformations (§II). So I deny that the way you are now is substantively psychologically like the way that you will be at every point in your endless future. But the way you are now might be psychologically continuous with the way you will be when glorified, and even with the way you will be at every point in your endless future. For the psychological transformations included in the hope of glory might occur gradually.

And there might be good reasons for concluding that the way you are now is psychologically continuous with the way you will be when glorified, or even with the way you will be at every point in your endless future. Suppose, for example, that an aspect of being glorified is being virtuous. Add that at least some virtues can be acquired only gradually, and in a way that preserves some sort of psychological continuity (see, e.g. Zagzebski, 1996, 116–25). Then conclude that the way you are now is psychologically continuous with the way you will be when glorified.

Or consider my claim that a person at a future time will have, at that time, what matters in survival for you only if you will be numerically identical with that person at that time (Ch. 3, §I). Add that you will be numerically identical with a person at a future time only if you are psychologically continuous with that person at that time. Then conclude that psychological continuity is necessary for survival. The hope of glory includes survival. So conclude that the hope of glory includes psychological continuity.

For what it is worth, I myself do not endorse either of these reasons (or any other reason) for concluding that the hope of glory includes psychological continuity. I have no opinion about the first reason, and I reject the second. I reject the second because I think that each of us is a human organism, and I deny that psychological continuity is necessary for any organism to endure (Ch. 3, §IV).

IV. Survival Does Not Come in Degrees

Consider:

> *The Why Question*: What way of being related to a (conscious) person at a future time explains why that person will have (at that time) what matters in survival for you?

Derek Parfit answers the Why Question in terms of psychological connectedness and/or psychological continuity (see Ch. 3, §II). Psychological connectedness comes in degrees. This leads Parfit to conclude that what matters in survival comes in degrees. Thus Parfit says:

> Even if [psychological] connectedness is not more important than [psychological] continuity, the fact that one of these is a relation of degree is enough to show that what matters in survival can have degrees.
>
> (1971, 21; see also 1984, 299)

Jeff McMahan's answer to the Why Question is a combination of physical and functional continuity and psychological continuity. McMahan claims that these continuities all come in degrees. And he takes this claim to imply that its being 'rationally appropriate' to have egoistic concern about oneself at a future time comes in degrees. Thus McMahan says:

> . . . there is a presumption that the degree of egoistic concern that a person is rationally warranted in having about his own future varies with the degree of functional continuity between himself now and himself in the future.
>
> (2002, 73)

And:

> I have argued that, in addition to physical and functional continuity in certain areas of the brain, psychological continuity is among the conditions of egoistic concern—not in the sense that it is *necessary* in order for egoistic concern to be rational, but in the sense that the *degree* to which it is rationally appropriate to be egoistically concerned about oneself in the future varies, other things being equal, with the degree to which one would be psychologically continuous with oneself in the future. (2002, 321)

McMahan says that its being 'rationally appropriate' to have egoistic concern about oneself in the future comes in degrees. So he would say

that its being appropriate to have future-directed self-interested concern with regard to one's future experiences comes in degrees. So he would say that what matters in survival comes in degrees.

Suppose that survival comes in degrees. That is, suppose that whether it is appropriate for you to first-personally anticipate, and have future-directed self-interested concern with regard to, a person's experiences at a future time is a matter of degree. Then you might think that the degree of appropriateness dwindles as the time until that person will have those experiences increases and as that person continues to change. Indeed—as we saw in Chapter 4 (§IV)—this is exactly what happens in David Lewis's (1976) story about Methuselah. So it is not surprising that Lewis joins Parfit and McMahan in thinking that survival comes in degrees, and in thinking this on the grounds that an answer to the Why Question comes in degrees.

Here is a story not about Methuselah, but about the hope of glory. It is now fully appropriate for you to have future-directed self-interested concern with regard to the experiences you will have on your first day of being (or starting to be) glorified. But as time passes, you will change. So it is now somewhat less appropriate for you to have future-directed self-interested concern with regard to the experiences you will have 30 years after being glorified. And it is now even less appropriate for you to have future-directed self-interested concern with regard to the experiences you will have 100 years after being glorified. And it is now not at all appropriate for you to have future-directed self-interested concern with regard to the experiences you will have at some point in the far distant future, such as, for example, in a billion years. A parallel point holds regarding first-personal anticipation. So eventually there will be no one who will have what matters in survival for you. So you will not survive forever. So the hope of glory is in vain.

If the story just told were accurate, then the hope of glory would be in vain. More generally, if this sort of story were accurate, then bona fide personal immortality would be impossible, unless, perhaps, one stays forever unchanged. But this story is not accurate. This is because—so I shall argue for the rest of this section—survival does not come in degrees.

Numerical identity does not come in degrees.[3] Numerical identity is a good answer to the Why Question (Ch. 2, §I). Moreover, every other good answer to the Why Question implies numerical identity (Ch. 3, §I). So every good answer to the Why Question is, or implies, an answer that does not come in degrees. Parfit, McMahan, and Lewis conclude that survival comes in degrees on the grounds that the answer to the Why Question comes in degrees. Similarly, I conclude that survival does not come in degrees on the grounds that every good answer to the Why Question is, or implies, an answer that does not come in degrees.

Let me elaborate. My defense of my answer to the Why Question began with the idea that it is appropriate for you now to have self-interested concern with regard to an experience that you are having right now. I claimed that such self-interested concern is appropriate because you are having that experience. In other words, this is appropriate because someone numerically identical with you is having that experience. So its being appropriate for you to have self-interested concern with regard to your current experiences is explained by your being numerically identical with the person having that experience (see Ch. 2, §I).

Its being appropriate for you to have self-interested concern with regard to the experiences you will have tomorrow has the same explanation. That is, this is explained by your being numerically identical with the person tomorrow who will then have those experiences (Ch. 2, §I). Numerical identity does not come in degrees. So that explanation does not hold to various degrees. So it is neither more nor less appropriate for you to have self-interested concern with regard to your current experiences than it is for you to have self-interested concern with regard to your experiences tomorrow.

Its being appropriate for you to have self-interested concern with regard to the experiences you will have tomorrow has the same explanation as does its being appropriate for you to have self-interested

[3] Even Parfit (1984, 298) says: 'Identity is logically one-one and all or nothing.' And Gareth Evans's (1978) and Nathan Salmon's (1981, 243) argument for the conclusion that identity cannot be indeterminate can be adapted to show that identity does not come in degrees. The first premise of this adaptation assumes for reductio that A is identical with B to some degree (other than 0 or 1).

concern with regard to the experiences you will have the day after tomorrow, or in 100 years, or in a billion years. Moreover, that explanation—your being numerically identical with the person who will have those experiences—does not come in degrees. So it is just as appropriate for you to have self-interested concern with regard to your experiences in a billion years as it is for you to have self-interested concern with regard to your current experiences.

The degree to which it is appropriate for you to have future-directed self-interested concern with regard to the experiences that you yourself will have does not decrease as time passes, or as you change, before having those experiences. Add that appropriate future-directed self-interested concern implies appropriate first-personal anticipation. Then conclude that the degree to which (a person who is identical with) you will have, at a future time, what matters in survival for you does not decrease as time passes, or as you change, between now and that future time (Ch. 1, §II). Moreover, only (a person who is identical with) you will have, at any future time, what matters in survival for you at that time (Ch. 3, §I). So survival does not come in degrees.[4]

V. The Tedium Objection

Bernard Williams (1973) claims that immortality would become unbearably tedious. That is, he claims that if you were to live forever, you would eventually be so bored that you would prefer ceasing to exist to continuing to exist. Suppose that Williams's claim is true. Then the hope of glory is not a hope for something that would be good for you. But the hope of glory is supposed to be a hope for something that would be good for you. Thus the Tedium Objection to the hope of glory.

[4] You survive as a person at a future time only if that person will be conscious at that time (Ch. 1, §III). I deny that it can be a matter of degree whether a person is conscious. But suppose I am wrong. Then it can be a matter of degree whether you survive as a person at a future time in virtue of its being a matter of degree whether that person is, at that time, conscious. This poses no threat to the hope of glory or to personal immortality more generally. For believers in personal immortality should deny that you will—simply as a result of time and change—eventually cease to be conscious.

Williams gives an argument for the claim that immortality would become unbearably tedious. That argument has premises like: our only reasons to live are (or perhaps come from) desires that we accomplish certain specific tasks. I am not going to focus on that argument. This is partly because I think that that argument is not compelling.[5] But it is mainly because that argument is not Williams's most influential defense of the claim that living forever would become unbearably tedious.

Williams's most influential defense of that claim is the idea that imagining what it would be like to live forever shows us that living forever would become unbearably tedious. At least arguably, this idea can be found at least as far back as Lucretius, who claims that Nature would rightly say:

> All things are always the same. If your body is not yet withered with age, nor your limbs decrepit with flagging, even so there is nothing new to look forward to—not though you should outlive all living creatures, or even though you should never die at all.
>
> (On the Nature of the Universe 3.946–3.950 [1994, 90])

This idea is also easy to find in contemporary philosophy. For example, Shelly Kagan says:

> Essentially, the problem with immortality seems to be one of inevitable boredom. The problem is tedium. You get tired of doing math after a while. After a hundred years, a thousand years, a million years, whatever it is, eventually you are going to say, 'Yes, here's a math problem I haven't solved before, but so what? I've just done *so much* math, it holds no appeal for me anymore.' Or, you go through all the great art museums . . . (2012, 243)

This idea is also central to some works of fiction, such as the short story 'The Dream' by Julian Barnes (1989) and (spoiler alert!) the closing episodes of the television series *The Good Place*.[6]

And Williams takes this idea to be supported by a work of fiction featuring Elina Makropulos and a magical elixir of youth. After living

[5] For critiques of that argument, see, e.g. Fischer (1994) and Rosati (2013).

[6] The relevant episodes of *The Good Place* allude to Williams (1973). And 'The Dream' might also have been influenced by Williams (1973), since the author's brother is the philosopher Jonathan Barnes.

as a young woman for over three hundred years, Makropulos decides to allow her life to come to an end. Williams (1973, 82) says that she makes this decision because: 'Her unending life has come to a state of boredom, indifference, and coldness. Everything is joyless: "in the end it is the same", she says, "singing and silence".'

Williams seems to take Makropulos's story to help us to imagine what it would be like for any of us to live forever. And Williams thinks that when you imagine this, you can see that you would eventually be so bored that you would prefer ceasing to exist to continuing to exist. But I reply that none of us can know what it would be like to be glorified and (say) a billion years old. So you cannot know that if you were glorified and a billion years old, you would be so bored that you would prefer ceasing to exist to continuing to exist.

A child cannot know what it would be like to have lived a rich and full life for seventy years. This is one reason that I think that you cannot now know what it would be like to have lived a rich and full life for a billion years. A second—related—reason is that being glorified and living for a billion years will bring about a change (or changes) in self. You cannot know what it would be like to have a (new and) different self (Ch. 4, §II). So you cannot know what it would be like to be glorified and a billion years old.

You cannot know what it would be like to be glorified and a billion years old. So you cannot know that you would be so bored that you would prefer ceasing to exist to continuing to exist. So Williams's most influential idea regarding immortality is false, at least when immortality includes glorification. That is, it is false that imagining what it would be like to live forever as glorified shows us that living forever as glorified would eventually become unbearably tedious.

Williams might agree that you cannot know what it would be like to live forever as glorified, or to live forever while undergoing other changes in self. So Williams might agree that immortality involving glorification or other changes in self is thereby immune to the Tedium Objection.[7] But he would still object to this sort of immortality. Williams would object that it is:

[7] Williams considers one other reply to the Tedium Objection: we find an activity so absorbing that we are not bored. He responds:

> . . . the attempt which some have made, to combine the best of continuous and of serial existence in a fantasy of very varied lives which are nevertheless cumulatively effective in memory. This might be called the *Teiresias* model. [In Greek mythology, Teiresias starts out as a man, is turned into a woman for seven years during which time she becomes a mother, and then is turned back into a man.] As that case singularly demonstrates, it has the quality of a fantasy, of emotional pressure trying to combine the uncombinable. One thing that the fantasy has to ignore is the connexion, both as cause and as consequence, between having one range of experiences rather than another, wishing to engage in one sort of thing rather than another, and having a character. Teiresias cannot have a character, either continuously through these proceedings, or cumulatively at the end (if there were to be an end) of them: he is not, eventually, a person, but a phenomenon. (1973, 94)

Williams does not claim that we can imagine what it would be like to have a series of 'very varied lives', and that imagining this shows us that this would become unbearably tedious. Instead, he seems to claim—put in my terms—that we cannot survive such a series. Put further in my terms, Williams seems to claim that we cannot survive changes in self.[8]

Kagan agrees with Williams:

> We can put the point in the form of a dilemma. Could immortality be something worth having forever? On the one hand, if we make the immortal person be *similar* to me, then boredom's going to set in . . . . On the other hand, if we solve the problem of boredom with progressive memory loss and radical personality changes, then maybe boredom wouldn't set in, but that life isn't anything I want for myself. It just doesn't *matter* to me that it's still going to be me, any more than it would matter to me if you were to just tell me, 'Oh, there'll be somebody *else* around . . . '. (2012, 245)

> But if one is totally and perpetually absorbed in such an activity, and loses oneself in it, then as those words suggest, we come back to the problem of satisfying the conditions that it should be me who lives forever, and that eternal life should be in prospect of some interest. (1973, 96)

This response equivocates on 'loses oneself'.

[8] Williams (1973) seems to deny that we can survive changes in self. This denial is inconsistent with the view favored in Williams (1970). For, as noted in Ch. 2 (§III), Williams (1970) favors the view that having the numerically same body delivers survival. And you can have the numerically same body as a person will have at a future time even if you do not have the same self as that person will have at that time.

But I reply to Williams and Kagan that we can survive changes in self (and we can survive forgetting). This includes not only the changes in self (and forgetting) that occur as a child becomes an adult (Ch. 4, §IV), but also the changes in self that occur as you become glorified and a billion years old (§II).

Perhaps there is a sense in which Teiresias does not continue to be the 'same person'. Perhaps there is a sense in which Teiresias exchanges one 'identity' for another, and then for yet another. And perhaps the hope of glory suggests that eventually each of us will, in some sense, become a 'different person' (maybe over and over again) or exchange one 'identity' for another (maybe ad infinitum). Chapters 5 (§VI) and 6 (§V) provide the resources to reconcile this with our surviving forever as glorified.

VI. The Irrationality Objection

I do not know what it is like to be a bat. So it is false that knowing what it is like to be a bat gives me a good reason to conclude that becoming a bat would not be good for me. Nevertheless, I think that I have good reasons to conclude that becoming a bat would not be good for me. This is because I know a lot about what bats are like and what they do. So I know that becoming a bat would deprive me of having rich interpersonal relationships and of doing philosophy. I deny that these—and so many other—deprivations would be compensated for, much less outweighed, by what I would gain by becoming a bat. So I conclude that becoming a bat would not be good for me.

You do not know what it would be like to be glorified and (for example) a billion years old. So it is false that knowing what this would be like gives you a good reason to conclude that being glorified and surviving forever would not be good for you (§V). Nevertheless, you might think that you have good reasons to conclude that being glorified and surviving forever would not be good for you. And you might say that this is just like my concluding that becoming a bat would not be good for me. But I disagree.

Suppose we know that the glorified will be perfectly loving and in communion with God and others who are perfectly loving. Even so, there is a lot that we do not know about what the glorified will be like

in (for example) a billion years, and a lot that we do not know about what they will then do. I think that because of this ignorance, it is not rational for any of us to conclude that the deprivations (if any) that come with being glorified and (for example) a billion years old would not be outweighed by the gains. So it is not rational for you to conclude that being glorified and surviving forever would not be good for you.

Samuel Scheffler (2016, 83–111) argues that living forever would not be good for you. His argument is mainly focused on living forever more or less unchanged, in more or less our current circumstances. This is not the hope of glory. But Scheffler does make some remarks that could be applied to the hope of glory. Scheffler says:

> If we never died, then we would not live lives structured by the kinds of values that now structure our own lives or by the kinds of values that have structured the lives of other human beings now and in the past. Moreover, it is at best unclear to what extent we would lead value-structured lives at all. What *is* clear, in any case, is that we would not live anything resembling what we now consider to be 'a life'. (2016, 207)

Let us suppose that to survive forever as glorified is not to live anything resembling what we now consider to be 'a life'. Scheffler might take this to constitute some sort of objection to the hope of glory (see, e.g. Scheffler, 2016, 98–9). But I say that this just reinforces my conclusion that none of us now knows enough about surviving forever as glorified, in communion with God and others, to rationally conclude that this would not be good for you.

But—you probably saw this coming—my conclusion leads to a new objection. According to the *Irrationality Objection*, none of us knows enough about surviving forever as glorified, in communion with God and others, to rationally conclude that this would be good for you.

Suppose this objection is right. This does not imply that surviving forever as glorified, in communion with God and others, would not be good for you. So the Irrationality Objection is not an objection to the hope of glory. Again, the target of the Irrationality Objection is not the hope of glory itself. Rather, the targets of this objection are those of us who believe that surviving forever as glorified, in communion with God and others, would be good for you. The Irrationality Objection says that it is not rational for us to believe this.

Here is my reply to the Irrationality Objection. We have learned by way of divine revelation that surviving forever as glorified, in communion with God and others, would be good for you. So it is rational for us to believe that surviving forever as glorified, in communion with God and others, would be good for you.

This reply to the Irrationality Objection will not convince those who deny that there is divine revelation, or who deny that such revelation could render rational a belief. But this reply should convince the targets of the Irrationality Objection. For the targets are those who already believe that surviving forever as glorified, in communion with God and others, would be good for you. And I assume that they believe this only because of revelation, revelation by way of (for example) scripture, religious tradition, or religious experience. So presumably they will accept that there is divine revelation and that such revelation can render rational a belief.

Those of us who share in the hope of glory might believe that it has been revealed that surviving forever as glorified would be good for everyone, and so good for you. Or we might believe that revelation about God's love, combined with revelation to the effect that we will survive forever as glorified, implies that surviving forever as glorified would be good for you. Or we might believe something else along these lines. But—whatever the details—those who share in the hope of glory have a reply to the Irrationality Objection: divine revelation renders rational the belief that it would be good for you to survive forever as glorified and in communion with God and others.

Suppose that revelation is a kind of divine testimony. Then my reply to the Irrationality Objection involves forming a belief about what would be good for us on the basis of divine moral testimony. Paddy McShane (2018) argues that the flourishing of intimate relationships is partly constituted by reliance on moral testimony. So perhaps a person's trusting God that glorification and endless survival, in communion with God and others, would be good for us partly constitutes the flourishing of that person's relationship with God.

I also think that trusting God in this way is an aspect of faith in God. That is, it is an aspect of faith in God to trust God that it will be good for us to be glorified and (for example) a billion years old and in communion with God and others, even though we do not know what it will be like to be glorified and a billion years old, and even though we

have few details about what we will then do and what features we will then have.

Compare: A six-year-old child realizes that she will grow up and no longer want to play with her favorite toy. If she is fairly precocious, she will also realize that not wanting to play with her favorite toy is just the tip of the iceberg: as an adult, she will have abandoned many of her current desires and values and projects. And if she is astonishingly precocious, this child might say that she is 'alienated' from her future adult self, which she takes to be 'an alien self' (cf. Paul, 2014, 125–31; 2015b).

So the child asks her parent whether being an adult will be good for her. Suppose the parent says 'yes'. Or suppose the parent is unusually candid and says that it will be good for her if she will do meaningful work, have supportive relationships, be free of constant pain, and so on. Suppose the child believes her parent. The parent's speaking honestly with the child, and the child's believing the parent, could partly constitute their having a flourishing parent–child relationship. And the child's believing her parent would be an aspect of her trusting that parent.

The preceding sections responded to objections to the hope of glory. Those objections could take aim at other versions of personal immortality. And other versions of personal immortality can avail themselves of the preceding sections' responses, or at least those other versions that, eventually, involve changes in self. But not so for this section's response to the Irrationality Objection. Consider, for example, personal immortality that has a naturalistic cause, such as a scientifically respectable elixir of youth. For all I say in this section, it might not be rational to believe that having this sort of immortality would be good for you, and this might not be rational for reasons akin to the Irrationality Objection: you have no idea what it would be like to be (for example) a billion years old, or what you would be doing then, or what sort of people—if any—you would be surrounded by, and so on.

VII. Conclusion

This chapter responded to various objections to the hope of glory, for the most part in ways that built on previous chapters. I think that

whether or not you share in the hope of glory, you should conclude that there is nothing incoherent or internally inconsistent or otherwise confused about the hope. So you can coherently hope for the good gift of being glorified and surviving forever, in communion with God and the best possible versions of those you love, and those you will come to love.

References

Barnes, Elizabeth (2015) 'What You Can Expect When You Don't Want to Be Expecting', *Philosophy and Phenomenological Research* 91: 775–86.

Barnes, Julian (1989) 'The Dream' in *A History of the World in 10½ Chapters* (New York: Alfred A. Knopf, Inc.).

Bradley, Ben (2009) *Well-Being and Death* (Oxford: Oxford University Press).

Broome, John (2013) 'The Badness of Death and the Goodness of Life' in Ben Bradley, Fred Feldman, and Jens Johansson (eds) *The Oxford Handbook of Philosophy of Death* (Oxford: Oxford University Press).

Butler, Joseph (1975 [1736]) "On Personal Identity" in John Perry (ed.) *Personal Identity* (Berkeley, CA: University of California Press).

Chadwick, Henry (1948) 'Origen, Celsus, and the Resurrection of the Body', *The Harvard Theological Review* 41: 83–102.

Chisholm, Roderick (1969) 'Reply' in Norman S. Care and Robert H. Grimm (eds) *Perception and Personal Identity: Proceedings of the 1967 Oberlin Colloquium in Philosophy* (Cleveland, OH: The Press of Case Western Reserve University).

Chisholm, Roderick (1976) *Person and Object: A Metaphysical Study* (La Salle, Ill.: Open Court Publishing Co.).

Cryanoski, David (2012) 'Neuroscience: The Mind Reader', *Nature* 486: 178–80.

DeGrazia, David (2005) *Human Identity and Bioethics* (Cambridge: Cambridge University Press).

Epicurus (1926) 'Letter to Menoeceus' in *Epicurus: The Extant Remains*, trans. Cyril B. Bailey (Oxford: Clarendon Press).

Evans, Gareth (1978) 'Can There Be Vague Objects?', *Analysis* 38: 208.

Feldman, Fred (1991) 'Some Puzzles about the Evil of Death', *Philosophical Review* 100: 205–27.

Fischer, John Martin (1994) 'Why Immortality Is Not So Bad', *International Journal of Philosophical Studies* 2: 257–70.

Geach, Peter (1980) *Reference and Generality: An Examination of Some Medieval and Modern Theories*, 3rd ed. (Ithaca, NY: Cornell University Press).

Haslanger, Sally (2003) 'Persistence through Change' in Michael J. Loux and Dean W. Zimmerman (eds) *The Oxford Handbook of Metaphysics* (Oxford: Oxford University Press).

Hawley, Katherine (2001) *How Things Persist* (Oxford: Clarendon Press).

Hawthorne, John (2003) 'Identity' in Michael J. Loux and Dean W. Zimmerman (eds) *The Oxford Handbook of Metaphysics* (Oxford: Oxford University Press).

Jackson, Frank (1982) 'Epiphenomenal Qualia', *Philosophical Quarterly* 32: 127–36.

James, William (1890) *The Principles of Psychology, Vol. 1* (New York: Henry Holt & Co.).

Johansson, Jens (2016) 'Animal Ethics' in Stephen Blatti and Paul F. Snowdon (eds) *Animalism: New Essays on Persons, Animals, and Identity* (Oxford: Oxford University Press).

Johnston, Mark (1992) 'Reasons and Reductionism', *Philosophical Review* 101: 589–618.

Johnston, Mark (1997) 'Human Concerns without Superlative Selves' in Jonathan Dancy (ed.) *Reading Parfit* (Oxford: Blackwell Publishers, Ltd).

Kagan, Shelly (2012) *Death* (New Haven, CT: Yale University Press).

Korsgaard, Christine (1989) 'Personal Identity and the Unity of Agency: A Kantian Response to Parfit', *Philosophy and Public Affairs* 18: 101–32.

Korsgaard, Christine (2009) *Self-Constitution: Agency, Identity, and Integrity* (Oxford: Oxford University Press).

Langsam, Harold (2001) 'Pain, Personal Identity, and the Deep Further Fact', *Erkenntnis* 54: 247–71.

Leibniz, G. W. (1989) Roger Ariew and Daniel Garber (eds) *Philosophical Essays* (Indianapolis, IN: Hackett Publishing Co.).

Leibniz, G. W. (1996) Peter Remnant and Jonathan Bennett (eds) *New Essays on Human Understanding* (Cambridge: Cambridge University Press).

Lewis, David (1976) 'Survival and Identity' in Amélie Oksenberg Rorty (ed.) *The Identities of Persons* (Berkeley, CA: University of California Press).

Locke, John (1975 [1689/1694]) Peter H. Nidditch (ed.) *An Essay Concerning Human Understanding* (Oxford: Clarendon Press).

Lucretius (1994) R. E. Latham (trans.) *On the Nature of the Universe* (New York: Penguin Putnam, Inc.).

MacIntyre, Alasdair (1984) *After Virtue*, 2nd ed. (Notre Dame, IN: University of Notre Dame Press).

McMahan, Jeff (2002) *The Ethics of Killing: Problems at the Margins of Life* (New York: Oxford University Press).

McShane, Paddy Jane (2018) 'The Non-Remedial Value of Dependence on Moral Testimony', *Philosophical Studies* 175: 629–47.

Merricks, Trenton (1994) 'Endurance and Indiscernibility', *Journal of Philosophy* 91: 165–84.

Merricks, Trenton (1995) 'On the Incompatibility of Enduring and Perduring Entities', *Mind* 104: 523–31.

Merricks, Trenton (1998) 'There Are No Criteria of Identity Over Time', *Noûs* 32: 106–24.

Merricks, Trenton (1999a) 'Endurance, Psychological Continuity, and the Importance of Personal Identity', *Philosophy and Phenomenological Research* 59: 983–97.

Merricks, Trenton (1999b) 'Persistence, Parts, and Presentism', *Noûs* 33: 421–38.

Merricks, Trenton (1999c) 'Review of Jonathan Dancy (ed.) *Reading Parfit*', *Philosophical Review* 108: 422–5.

Merricks, Trenton (2000) 'Perdurance and Psychological Continuity', *Philosophy and Phenomenological Research* 61: 195–7.

Merricks, Trenton (2001a) *Objects and Persons* (Oxford: Clarendon Press).

Merricks, Trenton (2001b) 'How to Live Forever without Saving Your Soul: Physicalism and Immortality' in Kevin Corcoran (ed.) *Soul, Body, and Survival* (Ithaca, NY: Cornell University Press).

Merricks, Trenton (2003) 'Review of Katherine Hawley, *How Things Persist*', *Mind* 11: 146–8.

Merricks, Trenton (2007) *Truth and Ontology* (Oxford: Clarendon Press).

Merricks, Trenton (2009) 'The Resurrection of the Body' in Thomas P. Flint and Michael C. Rea (eds) *The Oxford Handbook of Philosophical Theology* (Oxford: Oxford University Press).

Nagel, Thomas (1970) 'Death', *Noûs* 4: 73–80.

Nagel, Thomas (1986) *The View from Nowhere* (Oxford: Oxford University Press).

Olson, Eric T. (1997) *The Human Animal: Personal Identity without Psychology* (Oxford: Oxford University Press).

Olson, Eric T. (2003) 'An Argument for Animalism' in Raymond Martin and John Barresi (eds) *Personal Identity* (Oxford: Blackwell Publishing).

Parfit, Derek (1971) 'Personal Identity', *Philosophical Review* 80: 3–27.

Parfit, Derek (1984) *Reasons and Persons* (Oxford: Clarendon Press).

Parsons, Josh (2000) 'Must a Four-Dimensionalist Believe in Temporal Parts?', *The Monist* 83: 399–418.

Paul, L. A. (2014) *Transformative Experience* (Oxford: Oxford University Press).

Paul, L. A. (2015a) 'What You Can't Expect When You're Expecting', *Res Philosophica* 92: 1–23.

Paul, L. A. (2015b) 'Précis of *Transformative Experience*' and '*Transformative Experience*: Replies to Pettigrew, Barnes and Campbell', *Philosophy and Phenomenological Research* 91: 760–5 and 796–813.

Perkins, Franklin (2004) *Leibniz and China: A Commerce of Light* (Cambridge: Cambridge University Press).

Perry, John (1976) 'The Importance of Being Identical' in Amélie Oksenberg Rorty (ed.) *The Identities of Persons* (Berkeley, CA: University of California Press).

Ramelli, Ilaria L. E. (2019) *A Larger Hope? Universal Salvation from Christian Beginnings to Julian of Norwich* (Eugene, OR: Cascade Books).

Rea, Michael (1998) 'Temporal Parts Unmotivated', *Philosophical Review* 107: 225–60.

Rosati, Connie (2013) 'The Makropulos Case Revisited: Reflections on Immortality and Agency' in Ben Bradley, Fred Feldman, and Jens Johansson (eds) *The Oxford Handbook of Philosophy of Death* (Oxford: Oxford University Press).

Salmon, Nathan (1981) *Reference and Essence* (Princeton, NJ: Princeton University Press).

Schechtman, Marya (1996) *The Constitution of Selves* (Ithaca, NY: Cornell University Press).

Schechtman, Marya (2014) *Staying Alive* (Oxford: Oxford University Press).

Scheffler, Samuel (2016) *Death and the Afterlife* (Oxford: Oxford University Press).

Schroer, Jeanine Weekes and Schroer, Robert (2014) 'Getting the Story Right: A Reductionist Narrative Account of Personal Identity', *Philosophical Studies* 171: 445–69.

Shoemaker, David (2016) 'The Stony Metaphysical Heart of Animalism' in Stephen Blatti and Paul F. Snowdon (eds) *Animalism: New Essays on Persons, Animals, and Identity* (Oxford: Oxford University Press).

Shoemaker, Sydney (1985) 'Critical Notice of *Reasons and Persons*', *Mind* 95: 443–53.

Sider, Theodore (1996) 'All the World's a Stage', *Australasian Journal of Philosophy* 74: 433–53.

Sider, Theodore (2001) *Four-Dimensionalism* (Oxford: Oxford University Press).

Sidgwick, Henry (1907) *The Methods of Ethics*, 7th ed. (New York: Macmillan and Co. Ltd).

Smith, J. Warren (2018) 'Becoming Men, Not Stones: Epektasis in Gregory of Nyssa's Homilies on the Song of Songs' in Giulio Maspero, Miguel Brugarolas, and Ilaria Vigorelli (eds) *Canticum Canticorum: Analytical and Supporting Studies: Proceedings of the 13th International Colloquium on Gregory of Nyssa* (Leiden: Brill).

Strawson, Galen (2004) 'Against Narrativity', *Ratio* 17: 428–52.

Taylor, Charles (1989) *Sources of the Self* (Cambridge, MA: Harvard University Press).

Ullmann-Margalit, Edna (2006) 'Big Decisions: Opting, Converting, Drifting', *Royal Institute of Philosophy Supplement* 58: 157–72.

Velleman, J. David (1996) 'Self to Self', *Philosophical Review* 105: 39–76.

Velleman, J. David (2006) *Self to Self: Selected Essays* (Cambridge: Cambridge University Press).

Whiting, Jennifer (1986) 'Friends and Future Selves', *Philosophical Review* 95: 547–80.

Whiting, Jennifer (1991) 'Impersonal Friends', *The Monist* 74: 3–29.

Wiggins, David (1967) *Identity and Spatio-Temporal Continuity* (Oxford: Basil Blackwell).

Williams, Bernard (1970) 'The Self and the Future', *Philosophical Review* 79: 161–80.

Williams, Bernard (1973) 'The Makropulos Case: Reflections on the Tedium of Immortality' in *Problems of the Self* (Cambridge: Cambridge University Press).

Windt, Jennifer M., Nielsen, Tore, and Thompson, Evan (2016) 'Does Consciousness Disappear in Dreamless Sleep?', *Trends in Cognitive Sciences* 20: 871–82.

Wolf, Susan (1986) 'Self-Interest and Interest in Selves', *Ethics* 96: 704–20.

Zagzebski, Linda (1996) *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge* (Cambridge: Cambridge University Press).

Index

agency 26–27, 131–155
Agential Continuity Account of survival, defined 137
annihilation, badness of 109–110
appropriate, explained 7–14

Barnes, Elizabeth 110
Barnes, Jonathan 168n.6
Barnes, Julian 168
Bradley, Ben 109
Broome, John 109
Butler, Joseph 57

Chadwick, Henry 158n.1
change in self, defined 91
change in self, examples *see* growing up, transformations
Chisholm, Roderick 35n.3
criterion of personal identity over time 9n.1, 34n.2, 46–54, 69–83, 134–6, 143–144, 145n.5, 147, 153
Cryanoski, David 20n.6

death, badness of 109
Death Star 13, 14
DeGrazia, David 26, 28, 123

endurance, explained 33
epektasis 161–2
Epicurus 109n.7
Evans, Gareth 166n.3

Feldman, Fred 109
Fischer, John Martin 168n.5

Geach, Peter 79n.8
Gregory of Nyssa 161–2
growing up 103–107, 121–6, 127–8, 149, 162, 169, 174

Haslanger, Sally 39
Hawley, Katherine 43n.6, 44n.8, 72, 78
Hawthorne, John 79n.8
Hume, David 57

Jackson, Frank 107
James, William 158–9
Johansson, Jens 47, 49, 53–4
Johnston, Mark 30, 50n.11

Kagan, Shelly 28, 168, 170
Kid-Friendly Account of survival, defined 123
Korsgaard, Christine 30, 133–8, 146, 150–1

Langsam, Harold 35n.3
Leibniz, G. W. 7–9, 11, 12, 14, 15, 17, 23, 29, 54, 55, 57n.1, 87, 106, 109, 111, 152, 157, 158–9
Lewis, David 50, 73, 74, 106–7, 165–6
Locke, John 46–7, 57
Lucretius 109n.7, 168

MacIntyre, Alasdair 139, 140n.3
McMahan, Jeff 11, 17, 57n.1, 65, 67, 105n.4, 161–2, 164–6
McShane, Paddy Jane 173
Methuselah 106–7, 165

Nagel, Thomas 20, 109
narrative 4–5, 85, 88, 89, 101, 113–133, 139–155, 160–1
Narrative Continuity Account of survival, defined 139
Nielsen, Tore 21n.8

Olson, Eric T. 10, 11, 17, 18, 48, 49, 53–4, 76, 78
Origen 158

Parfit, Derek 4, 17, 20n.5, 25–6, 47, 52, 57, 59n.2, 65–83, 84, 86, 134, 138, 141, 152, 157, 164–166

Parsons, Josh 36n.4
Paul, the Apostle 157
Paul, L. A. 92, 94, 98–9, 157, 174
perdurance, explained 36
Perkins, Franklin 7
Perry, John 87–93, 100–102, 107

Ramelli, Ilaria L. E. 158n.1
Rea, Michael 40
resurrection
—as hope for eternal life 160n.2
—as way of becoming spherical 158
Rosati, Connie 168n.5

Salmon, Nathan 166n.3
Schechtman, Marya 10–12, 17, 18, 29, 57n.1, 88–93, 100–102, 113–18, 122–5, 129
Scheffler, Samuel 172
Schroer, Jeanine Weekes 140–1
Schroer, Robert 140–1
Self-Narrative Account of survival, defined 115
Selfer, defined 91–92
Shoemaker, David 18, 47, 49, 53–4
Shoemaker, Sydney 17, 65, 67, 68
Sider, Theodore 43n.6, 44, 72, 78
Sidgwick, Henry 57
Smith, J. Warren 161
stage theory, explained 43–44, 78
Strawson, Galen 119

Taylor, Charles 129, 140n.3
Thompson, Evan 21n.8
transformations 92–3, 98–9, 107–111, 118–9, 126–128, 145–150, 152–154, 160–3, 169–171
see also growing up

Ullmann-Margalit, Edna 94

Velleman, J. David 10–12, 17, 21, 30, 54, 94–101, 104, 107

Whiting, Jennifer 10–12, 17, 29, 54, 89–93, 100–101, 102
Wiggins, David 72
Williams, Bernard 47–8, 167–171
Windt, Jennifer M. 21n.8
what matters in survival
—for the nature of what matters in survival, *see* the What Question
—for what delivers what matters in survival, *see* the Why Question
What Question 1, 7, 14, 16, 18, 22, 26, 28, 34, 42, 63, 65, 90, 115, 136
—my answer, defended 14–20
—other answers 16–18
Why Question 1, 19, 22, 28, 29, 31–35, 45, 48, 54, 55, 59, 67, 83, 85, 101, 111, 113, 115, 120, 130n.1, 131, 133, 137, 139, 142, 151, 155, 159, 162, 164
—my answer, defended 31–35, 151–5
—other answers 2, 46–54, 67–69, 83–85, 101–103, 115–118, 123–6, 133–7, 139–142, 164–5
—why every good answer implies numerical identity 57–64
Wordier Why Question 1, 11, 19, 35
—my answer, defended 31–35
—other answers 11
Wolf, Susan 26n.9

Zagzebski, Linda 163